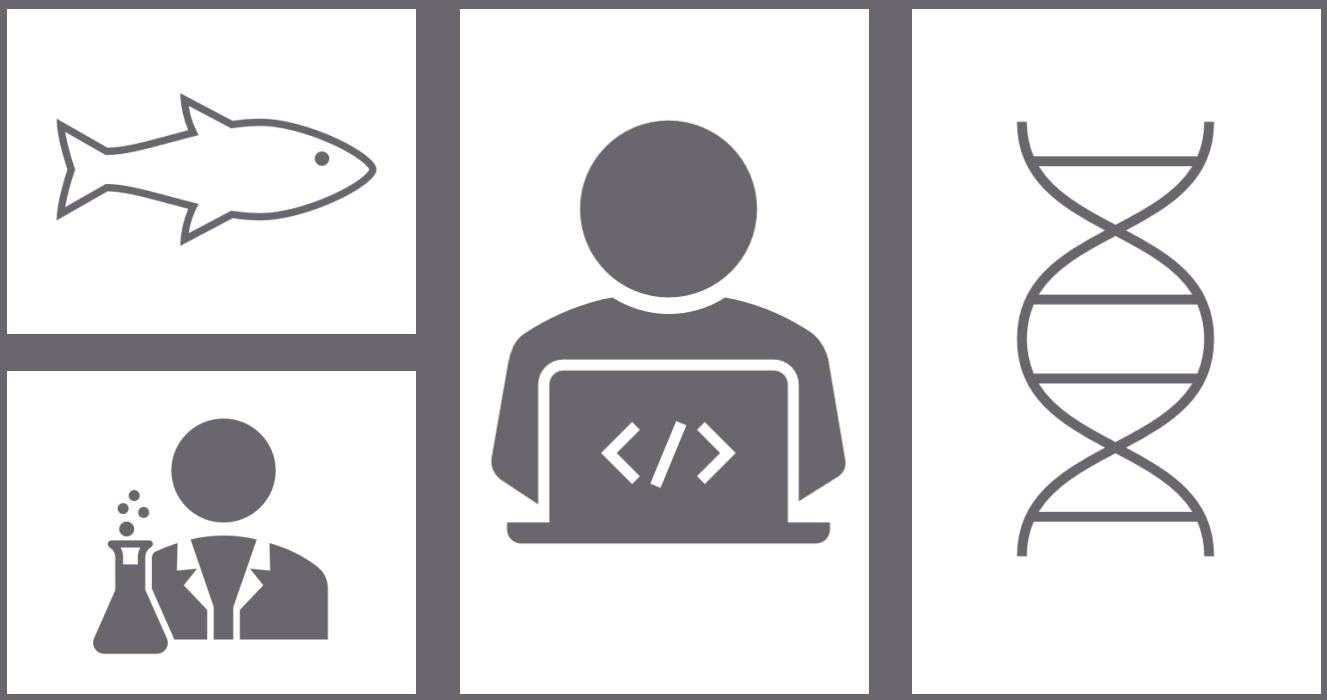
Transcriptomic landscape of pancreatic cell types and the identification of a common Pax6-dependent gene regulatory network in pancreatic and intestinal endocrine cells

ARNAUD LAVERGNE

A thesis manuscript presented for the degree of Doctor of Philosophy in Sciences



Supervised by: **Dr. Bernard Peers**

Laboratory of Zebrafish Development and Disease Models

GIGA Institute

Liège University, BE

July 2020

LIÈGE université GIGA institute

Zebrafish Development & Disease Models

fnrs LA LIBERTÉ DE CHERCHER

FONDS LÉON FREDERICQ

Transcriptomic landscape of pancreatic cell types and the identification of a common Pax6-dependent gene regulatory network in pancreatic and intestinal endocrine cells

Arnaud Lavergne

A thesis manuscript presented for the degree of
Doctor of Philosophy in Sciences

Supervised by:
Bernard Peers, PhD
Laboratory of Zebrafish Development and Disease Models

GIGA Institute
Liège University, BE
July 2020

Acknowledgements

A PhD thesis is quite a rewarding journey. To reach the end of this long trip it will take a lot of energy, but it is worth it. At times, you need self-confidence and perseverance. However, if you surround yourself with the right people it will be an easy ride. I would like to thank all those who have been by my side during these last years and who have helped me achieve my goals. Firstly, I would like to thank those who made me grow as a young scientist. The jury members, Dr. Guy Rutter, Dr. Patrick Callaerts, Dr. Carole Charlier, Dr. Patrick Meyer, and chairman Dr. Denis Baurain, to have accepted to review this manuscript and to be part of the jury. I would also like to thank the members of my thesis committee: Dr. Isabelle Manfroid, Dr. Carole Charlier, Dr. Patrick Meyer, and Dr Benoit Charloteaux, who followed the progress of my project over the last years. I am grateful to my supervisor Dr. Bernard Peers who gave me the opportunity to pursue a research project during my PhD. Together with Dr. Marianne Voz and Dr. Isabelle Manfroid, they have accepted me in their laboratory since my master thesis. They gave me the tools and the required environment to develop my skills and they trusted me to manage the bioinformatics analysis of their data.

Je voudrais remercier plus particulièrement Bernard Peers de m'avoir guidé et encadré tout au long de cette thèse. Tu m'as laissé la liberté nécessaire pour me permettre de prendre des initiatives et d'apprendre les compétences que je désirais, tout en sachant me canaliser quand il le fallait. J'ai appris beaucoup sous ta supervision et tu as su me transmettre ton optimisme et ton goût pour la science. Malgré toute ton expérience, tu gardes une saine curiosité et une passion pour ce que tu fais : je me souviens d'un jour où pendant une séance de dissection (chronométrée), tu avais dévié du pancréas pour aller voir « un signal intéressant au niveau de l'œil ». Merci de m'avoir motivé, encouragé et soutenu pour m'amener au bout de ce projet. J'ai toujours pu compter sur toi pendant ma thèse, et je te remercie sincèrement de la confiance que tu m'auras accordé tout au long de ces années.

Je remercie également Isabelle et Marianne pour le soutien qu'elles m'ont apporté. Vous formez avec Bernard un groupe de choc à la tête de ce labo. Isabelle, tu as été la première à m'accueillir dans ton groupe lors d'un stage où j'avais pu découvrir le monde fascinant de la régénération. Je te remercie également pour toute l'aide que tu m'as apporté tout au long de ma thèse, mais également pour ta disponibilité et ta confiance pour l'analyse de tes données. Marianne, merci pour ton enthousiasme et ton aide dans ce projet. Nos efforts combinés ont permis d'aboutir à la rédaction d'une chouette publication. Merci également de ta confiance pour l'analyse de tes données.

Un tout grand merci à Tefa. Tu m'as ouvert la voie de la transcriptomique au niveau du pancréas chez le Zebrafish. Ton optimisme et ta persévérance m'ont servi d'exemple à suivre. Je suis heureux d'avoir partagé avec toi ces deux sujets de recherche enrichissant et je te souhaite beaucoup de réussite dans ton parcours scientifique.

Dede, comment ne pas te réserver un paragraphe rien que pour toi. Je te remercie d'avoir été un vrai camarade au et en dehors du labo. J'ai énormément de bons souvenirs partagés avec le « BG », notamment dans le bureau qu'on s'était aménagé. Même si aucun de nous n'aura réussi son insertion CRISPR, on aura cependant réussi bien d'autres choses (peut-

être moins sérieuses) : les excursions dans le CHU, les batailles d'Hedgewars, les « but alors you are french », les « lekker by me, lekker by me », les « Fatal et Vito » ou encore les drakkars vikings, bref on n'a pas chômé. C'est un vrai plaisir de t'avoir comme ami, et comme l'a dit l'autre : « Que la force soit avec toi, Harry » - Gandalf.

Thom Poll, Amandine et Marie, vous aussi notre amitié aura su s'étendre au-delà du labo et m'aura permis d'arriver au bout de ce projet. Thom, on aura fait un bout de chemin ensemble depuis le master, et notre amitié n'aura fait que se renforcer tout au long. Nos multiples séjours sur les pistes enneigées t'auront valu le titre d' « Artiste ». Tu es un vrai moteur pour notre petit groupe grâce à ton énergie inépuisable. Nos pauses Roland Garros me manquent, mais j'espère qu'on pourra prochainement pratiquer plutôt que de regarder. Amandine, ma sista, la bestie de ma bestie. Ta bonne humeur et ton énergie ont apporté beaucoup dans le labo. Tu es une personne pétillante qui possède un très bon sens de l'humour. Nos goûts culinaires et nos caractères nous ont permis de lier une chouette amitié même si je redoute toujours de me retrouver dans ton équipe lors de n'importe quelle compétition. Plus sérieusement, je te remercie sincèrement pour ton soutien durant ces derniers mois de thèse. Marie, la petite dernière arrivée. Je n'étais déjà plus officiellement au labo à ton arrivée, mais pourtant j'ai vite accroché à ta personnalité. Tu es quelqu'un d'une grande gentillesse avec qui j'ai très vite apprécié discuter. Tes origines parisiennes n'ont pas été un obstacle à notre amitié et je pense qu'avec le groupe, on aura encore l'occasion de faire de chouettes activités prochainement.

Jojo, Vivi et Alice, cela aura été un plaisir de partager un « U » avec vous pendant ma thèse. Votre bonne humeur, votre énergie et le son de la radio auront vraiment rendu les manips plus agréables chaque jour. Je vous souhaite à chacune le meilleur pour la suite.

Claudia et Ana, l'exotisme du labo. Cela a été un plaisir de vous rencontrer et de partager avec vous des moments dans le consortium ZENCODE. Claudia, j'ai appris à connaître ton caractère catalan et ta cool-attitude. Tu es une personne « buena onda » à qui je souhaite beaucoup de bonnes choses. Ana, tu es une personne très talentueuse et je te souhaite de la réussite dans ton parcours professionnel.

Je remercie également les personnes que j'ai pu rencontrer au labo pendant ces quelques années : Dave, Tom Wind, Marine, Stella, Justine, Claudio, Laura, Maurijn, Renaud, Oli, … Chacun de vous a contribué à créer une bonne ambiance et un environnement de travail plaisant.

Je voudrais également remercier Alice et Benoit (et l'équipe BIF), qui au cours des deux dernières années m'ont permis de grandir et de passer un nouveau cap dans mon développement professionnel. Vous m'avez permis de retrouver un nouvel environnement tout aussi plaisant et je vous remercie pour votre soutien et vos encouragements pour que je termine cette thèse.

Mes remerciements vont également aux personnes travaillant dans l'animalerie Zebrafish ainsi que dans les différentes plateformes du GIGA : imagerie, cytométrie en flux et génomique. Je remercie notamment mes nouveaux collègues de la plateforme génomique : Wouter, Lati, Manon et Emilie.

Anne-Sophie, tu es celle qui a changé la donne. Tu as été mon boost, ma motivation et ma force pour arriver au bout de ce projet. Heureusement, ce n'est pas la fin de notre projet

à nous et notre duo est maintenant lancé pour aller bien plus loin. Je n'en rajoute pas d'avantage, car j'ai la chance de pouvoir te dire tout le bien que je pense de toi chaque jour. Merci pour tout Nam !

Je voudrais également remercier mes amis. Tous ceux qui m'ont soutenu en dehors du labo et qui pendant ces années de thèse, m'ont permis de décompresser et de prendre du recul quand il le fallait. Marcia et Antoine, vous êtes ceux qui m'ont permis de devenir celui que je suis. Je vous remercie d'avoir été à mes côtes depuis maintenant plus de 15 ans et pour tous les bons moments qu'on a partagés. Ingrid et Damien, vous avez naturellement rejoint notre groupe et je vous remercie pour votre amitié ces dernières années.

Je tiens finalement à remercier ma famille qui m'a soutenu et permis de réaliser ses études. Laurent, je te remercie pour toutes nos discussions, tous les moments qu'on partage encore aujourd'hui et pour tous ce qu'on aura traversé. Oli, merci d'avoir toujours répondu présent quand j'en ai eu le besoin. Je ne te l'ai pas souvent dit, mais tu as toujours été un modèle de détermination et de travail dont j'ai essayé de m'inspirer pendant cette thèse.

Je remercie enfin ma maman qui m'a permis de réaliser les études que je voulais. Tu m'as fait confiance et tu m'as soutenu même quand tu ne comprenais pas toujours mes décisions. Cette thèse, c'est un peu l'aboutissement de tout ce que tu m'as donné pour me permettre d'atteindre mes objectifs. Merci maman.

Table of contents

Preface

This manuscript is submitted for obtaining the degree of Doctor of Philosophy in Sciences. The work described here includes mainly the presentation of bioinformatics analyses of high throughput sequencing (HTS) data. These include RNA-seq, ChIP-seq and scRNA-seq data that have been generated in the laboratory as well as data that were publicly available. The results are organized around two main topics:

The identification of conserved transcriptomic signatures for all major pancreatic cell types. This work has already been associated to a first publication. I present in this manuscript our efforts to complete the previous results using the emerging scRNA-seq data.
Credit of the work: The zebrafish RNA-Seq datasets were prepared by Estefania Tarifeño-Saldivia, Isabelle Manfroid and David Bergemann. Estefania Tarifeño-Saldivia performed the bioinformatic analysis of these data. I performed the analysis of the publicly available mammal datasets and the comparison with the zebrafish results. Alice Bernard contributed to the project by generating *myt1* and *cdx4* zebrafish mutants. *In situ* hybridization were performed under the collective effort of Estefania Tarifeño-Saldivia, Alice Bernard, Keerthana Padamata and myself.

The role of Pax6b during pancreatic and intestinal endocrine cell development in zebrafish. We analyzed the transcriptional changes in pancreatic endocrine cells and enteroendocrine cells due to a *null* mutation in the *pax6b* gene. This reveals the pax6-dependent gene regulatory network in both tissues. These results have been included in a manuscript submitted to **BMC Biology** and it has currently been **re-submitted after a second round of minor reviewing**. I also present work concerning the identification of Pax6 genomic binding sites using ChIP-seq data in the laboratory.
Credit of the work: The study and the characterization of enteroendocrine cells have been conducted by Justine Pirson, Anne-Sophie Reuter and Mar-

ianne Voz. They performed the *in situ* hybridization experiments in EECs. Estefania Tarifeño-Saldivia and Justine Pirson carried out the RNA-seq experiments for WT and *pax6b*$^{-/-}$ mutant PECs and EECs. I performed all the bioinformatic analyses concerning the transcriptomic comparisons, the differential gene expression analysis and the identification of the cell subtype markers. I performed the ISH experiments in PECs and EECs of the *pax6b* mutants. I carried out all the process for the ChIP-seq experiments: collecting the embryos, performing chromatin immunoprecipitation, preparing the ChIP-seq libraries, and performing the analysis.

Contributions by the other members of the laboratory are again described in the corresponding sections.

Abbreviations

bHLH	**b**asic **H**elix **L**oop **H**elix
CNE	**C**onserved **N**on-genic **E**lement
CNS	**C**entral **N**ervous **S**ystem
DE	**D**ifferential **E**xpression / **D**ifferentially **E**xpressed
DM	**D**iabetes **M**ellitus
DPF	**D**ays **P**ost **F**ertilization
DR	**D**anio **R**erio
EEC	**E**ntero**E**ndocrine **C**ell
ENS	**E**nteric **N**ervous **S**ystem
FACS	**F**luorescent-**A**ctivated **C**ell **S**orting
GFP	**G**reen **F**luorescent **P**rotein
GI	**G**astro**I**ntestinal
GPCR	**G**-**P**rotein **C**oupled **R**eceptor
GSIS	**G**lucose **S**timulated **I**nsulin **S**ecretion
GRN	**G**ene **R**egulatory **N**etwork
HD	**H**omeo**d**omain
HS	**H**omo **S**apiens
HPF	**H**ours **P**ost **F**ertilization
HTS	**H**igh-**T**hroughput **S**equencing
IMGU	**I**nsulin-**M**ediated **G**lucose **U**ptake
MM	**M**us **M**usculus
MODY	**M**aturity **O**nset **D**iabetes of the **Y**oung
MYA	**M**illion **Y**ears **A**go
NIMGU	**N**on-**I**nsulin-**M**ediated **G**lucose **U**ptake
NGS	**N**ext **G**eneration **S**equencing
PCA	**P**rincipal **C**omponent **A**nalysis
PD	**P**aired **D**omain
PEC	**P**ancreatic **E**ndocrine **C**ell
T1D	**T**ype **1** **D**iabetes

T2D	**T**ype **2** **D**iabetes
TGD	**T**eleost **G**enome **D**uplication
TSS	**T**ranscription **S**tart **S**ite
UMAP	**U**niform-**M**anifold **A**pproximation for **P**rojection for dimension reduction
UMI	**U**nique **M**olecular **I**dentifier
WGD	**W**hole **G**enome **D**uplication

Introduction

1 General Background

1.1 Overview of Digestive System

Digestive system regroups all organs playing role in the processus of digestion (Figure 1) . Gastrointestinal (**GI**) tract is the main performer and its intervention starts directly after food intake. Its functions are as varied as the mechanical food integrity disruption, the progression of food bolus, the extraction and the absorption of nutrients or the elimination of remaining solid wastes in the form of feces. Mobility of food bolus inside the gut is performed through a large smooth muscle tissue maintaining constant tonus in semi-tense state. It can be locally adapted to allow food to progress (**peristalsis**) or to be stored in some specific portions [1]. Each region of the GI tract produces specific secretions containing enzymes which will contribute to digest food. They are joined in this function by accessory organs of digestion localized along the gut. Accessory organs include both the tongue and the salivary glands, the pancreas, the liver and the gallbladder. The first two are involved in tasting and they respond by secreting mucous, to facilitate swallowing, and digestive enzymes for breaking food. The pancreatic juice and the bile, produced by liver and gallbladder, react with the *chyme* and create environment which facilitates food degradation by the pancreatic digestive enzymes and the other GI enzymes.

The global process of digestion needs to be strongly regulated and controlled in order to work optimally. This is why the digestive system shelters an intrinsic nervous system together with a specific part of the endocrine system. Enteroendocrine cells (**EECs**) scattered in the GI tract represent the largest endocrine organ in the body in terms of number of cells. Pancreatic endocrine cells (**PECs**) form another endocrine tissue inside the digestive system which plays key role in glucose homeostasis by controlling the balance between *insulin/glucagon* release. PECs are grouped in clusters, called the islets (**Islets of Langerhans**), which are scattered in pancreatic tissue. This dispersion of both EECs and PECs inside digestive organs suggests an interaction between exocrine and

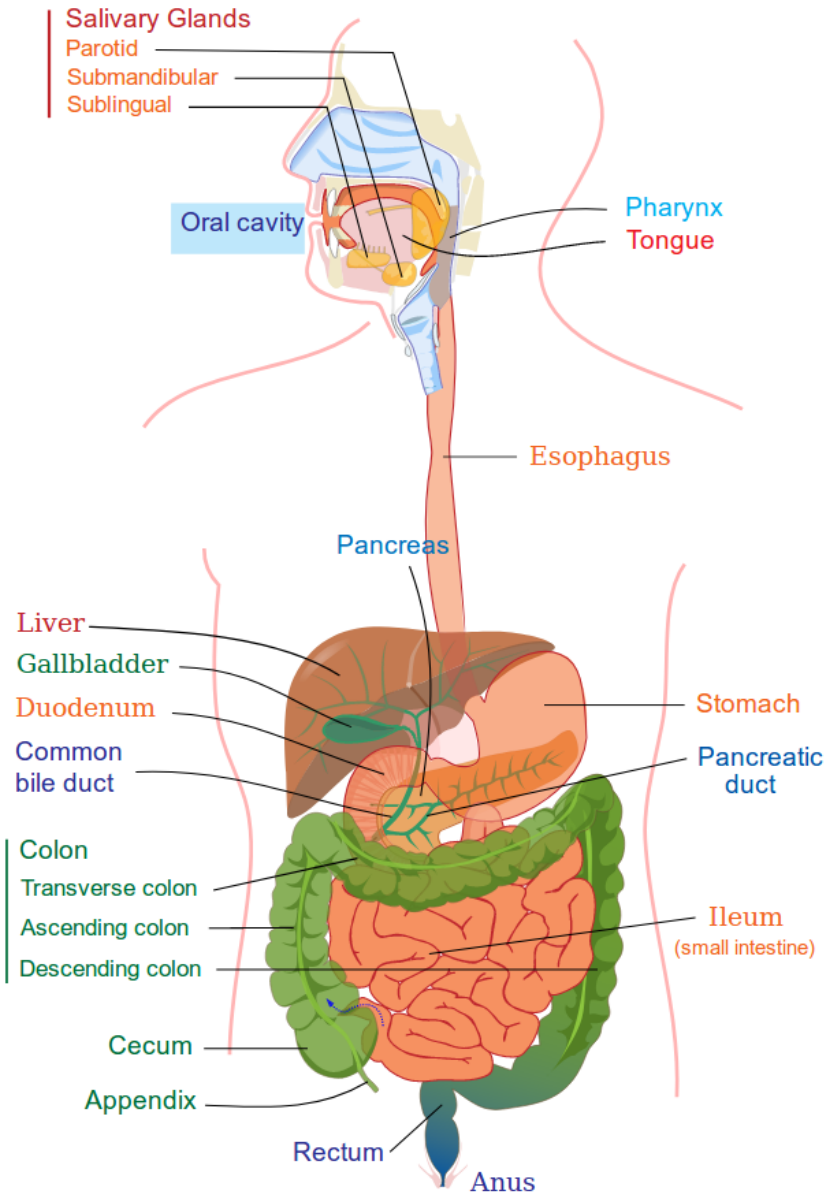


Figure 1: **General overview of the gastrointestinal tract**. Upper GI tract extends from oral cavity to the duodenum. Lower GI tract extends from the end of the duodenum to the anus. The tongue, the salivary glands, the pancreas, the liver and the gallbladder constitute the accessory organs of digestion (*https://en.wikipedia.org/wiki/Gastrointestinal_tract*).

endocrine cells. Endocrine cells act as sensors and initiate optimally the local response. They also inform the rest of the body about what is happening in specific regions during bolus progress. Finally, endocrine cells are closely associated with the enteric nervous system (ENS) and the central nervous system (CNS) for controlling global GI processes.

1.2 Gastrointestinal tract

1.2.1 Anatomical description and functional organization

GI tract, or digestive tract, consists of all organs crossed by food intake during the digestive process. In humans, it starts with the mouth and it runs successively through pharynx, oesophagus, stomach, all parts of the intestine (duodenum, jejunum, ileum, cecum, colon, rectum) and ends with the anus. GI tract can be divided in two spatial regions separated by suspensory muscle of duodenum: upper and lower gastrointestinal tracts.

1.2.2 Upper Gastrointestinal tract

Upper GI consists of all organs until the duodenum included [2] and it represents portions where all secretions are continually added to food in order to break it down into small nutrients before they can be absorbed. Salivary glands produce the first secretions to contribute to the first bolus. They are represented by the major salivatory glands which include parotid, submandibular and sublingual salivatory glands and by many minor salivatory glands distributed in buccal cavity [3] (Figure 2). There are two different types of secretions: serous and mucous. Serous secretions contains mainly two digestive enzymes, *amylase* and *lipase*, which initiate digestion of starch and lipids respectively [3, 4]. Mucous secretions contain *mucin* which acts as lubricant to facilitate mechanical processes like mastication, swallowing and the bolus progress through oesophagus. The junction between the oesophagus and the stomach, called the *cardia*, is kept shut by muscles of the oesophagus and by the diaphragm. These muscles will relax during swallowing to let food enters stomach [5].

The stomach consists of different parts: the *fundus* is the upper-part, the *body* (or *corpus*) and the *antrum* represent the main region, and the *pylorus* is the final portion connected

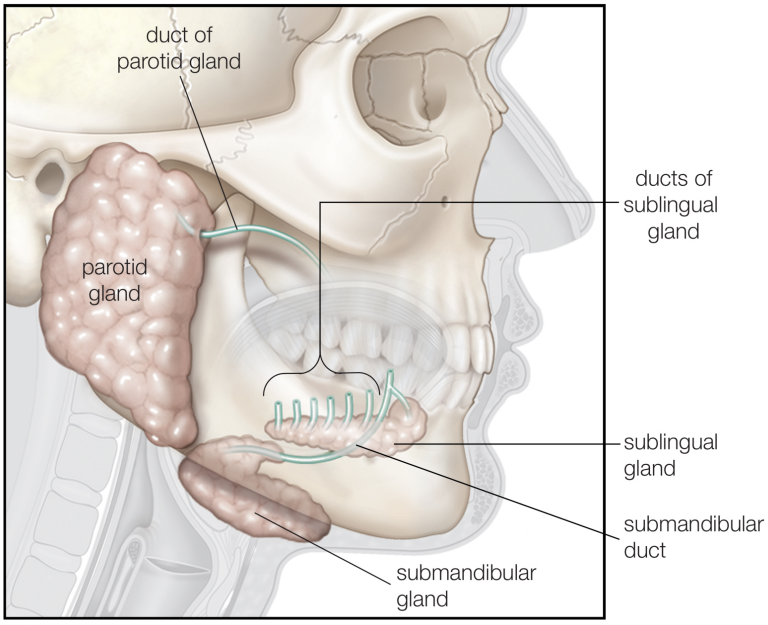
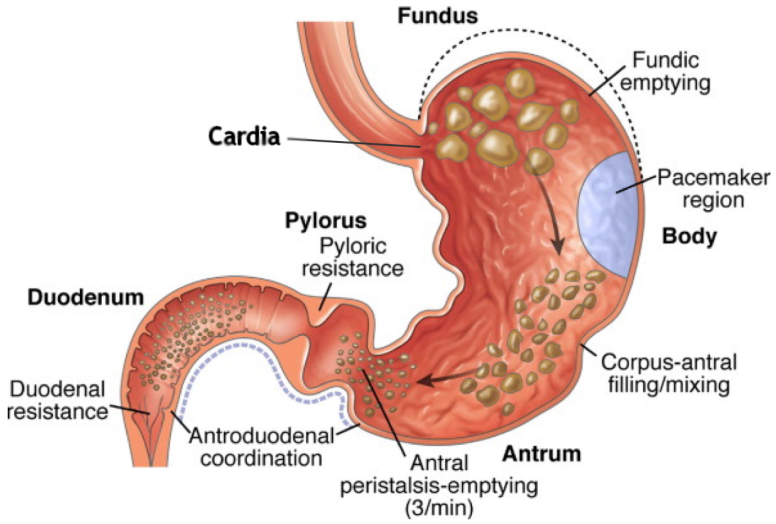


Figure 2: **The salivary glands**. Major salivatory glands include parotid, submandibular and sublingual salivatory glands. Multiple ducts lead their secretion in the buccal cavity.



(adapted from https://www.memorangapp.com/flashcards/249956/Gastric+Motility+and+Secretions)

Figure 3: **The anatomy of the stomach**. Food bolus enters the stomach cavity from the oesophagus and reach the duodenum through the pyloric sphincter. Different regions are observed: the fundus, the body, the antrum and the pylorus

to the duodenum through pyloric canal and sphincter [5]. Stomach serves as a storage point where gastric secretions create an acid environment which will bathe bolus to progressively transform it into the *chyme* [6, 7]. Gastric smooth muscles perform regular contractions to mix and to trigger gastric emptying. *Pepsinogen* is the proenzyme secreted in the stomach and the acidification of the gastric content is mandatory for its conversion to *pepsin*. *Pepsin* initiates protein degradation in order to reduce the size of nutrients [7]. Once it is sufficiently digested, *chyme* can run through the pyloric sphincter to reach the duodenum (Figure 3). Duodenum forms the smallest part of intestine but plays a key role for the absorption in the following intestine sections. Descending part of the duodenum receives terminations of **Wirsung duct** (pancreatic juice) and **bile duct** (bile from liver and gallbladder). Pancreatic juice has a high concentration of bicarbonates to balance acidity of *chyme* inherited from stomach and contains pancreatic digestive enzymes required for the rest of digestion [8]. Bile has hydrophillic and hydrophobic properties which plays role in the emulsification of lipids. This leads to the essential formation of fatty acid micelles to allow their absorption in intestine.

1.2.3 Lower Gastrointestinal tract

Lower GI consists of the rest of small intestine (jejunum and ileum) and the large intestine (colon) [2]. Its main roles are the absorption of all nutrients together with the disposal of solid waste in the form of feces. In both jejunum and ileum, mucosal layer increases the surface in order to absorb maximum of nutrients. It is achieved by formation of different macro and micro-anatomical structures. **Circular folds** are large projections of the mucosal layer into the lumen which slows down the progress of *chyme*. Both **villi** and **crypts** (crypt of Lieberkühn) form an intermediate structural organization of the intestinal epithelium, closely associated with its stratification (Figure 4). Indeed, the proliferative stem cells, found at the bottom of crypts, differentiate progressively into new enterocytes and secretory cells as they migrate along villi to balance the loss of senescent cells at the apex [9]. Intestinal epithelium cells have the highest turnover rate in adult mammals body and almost complete renewal of enterocytes occurs every 2 to 6 days [10] (Figure 5). **Microvilli** form micro-structures in comb shape adopted by enterocytes at their apex which highly increase the cell surface in contact with lumen

in order to maximize the absorption. Both jejunum and ileum are anatomically closed but they present slightly functional differences. The jejunum is where the majority of nutrients are absorbed while the ileum is responsible for absorption of vitamin B12, bile acids along with the remaining nutrients [11].

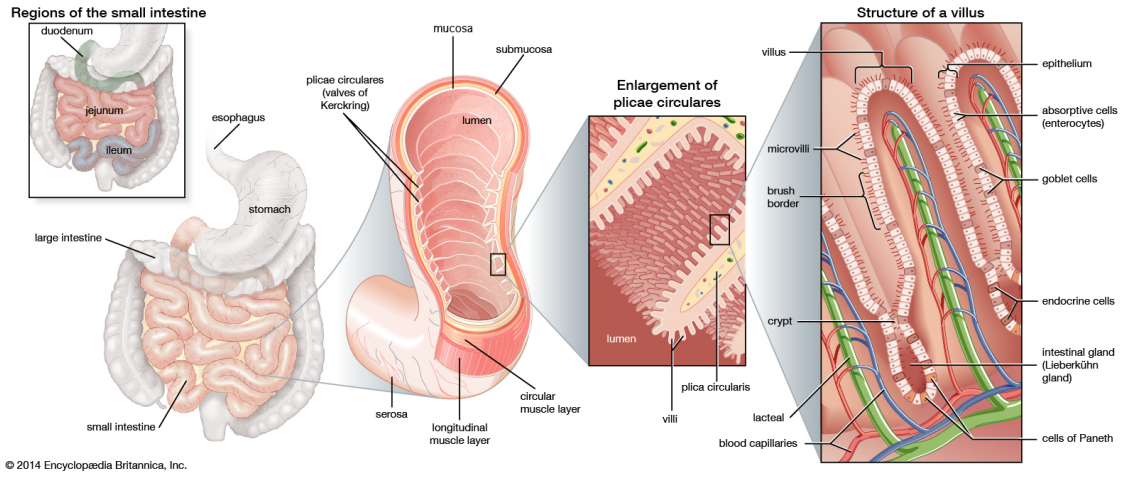


Figure 4: **The anatomical organization of the intestin**. The macro and micro-anatomical structures adopted by the mucosal layer increase greatly the surface of absorption in the intestin.

Colon does not possess the absorptive structural properties of the small intestine as everything that reaches the colon is intended to be evacuated. The colon reabsorbs remaining water molecules and causes waste desiccation to constitute *feces* which will be evacuated through the anus sphincter [13, 14]. In the same time, microbiota starts fermentation of products inside the colon to extract the remaining nutrients and it produces additional waste important for *feces* formation [15]. Strong smooth muscles contract sequentially (by waves) to assure peristalsis and the solid waste mashing in order to facilitate formation of *feces*. Mobility in the colon is slower than in the small intestine but waves allow progress of *feces* until rectum [14].

1.2.4 Nervous control of Gastrointestinal tract

Digestive tract needs to be well controlled to work optimally. It differs from all peripheral organs by sheltering an extensive intrinsic nervous system, the enteric nervous system (ENS). The ENS originates from the neural crest with enteric neural crest-derived cells colonizing the entire GI tract. The ENS develops to constitute a network of neurons

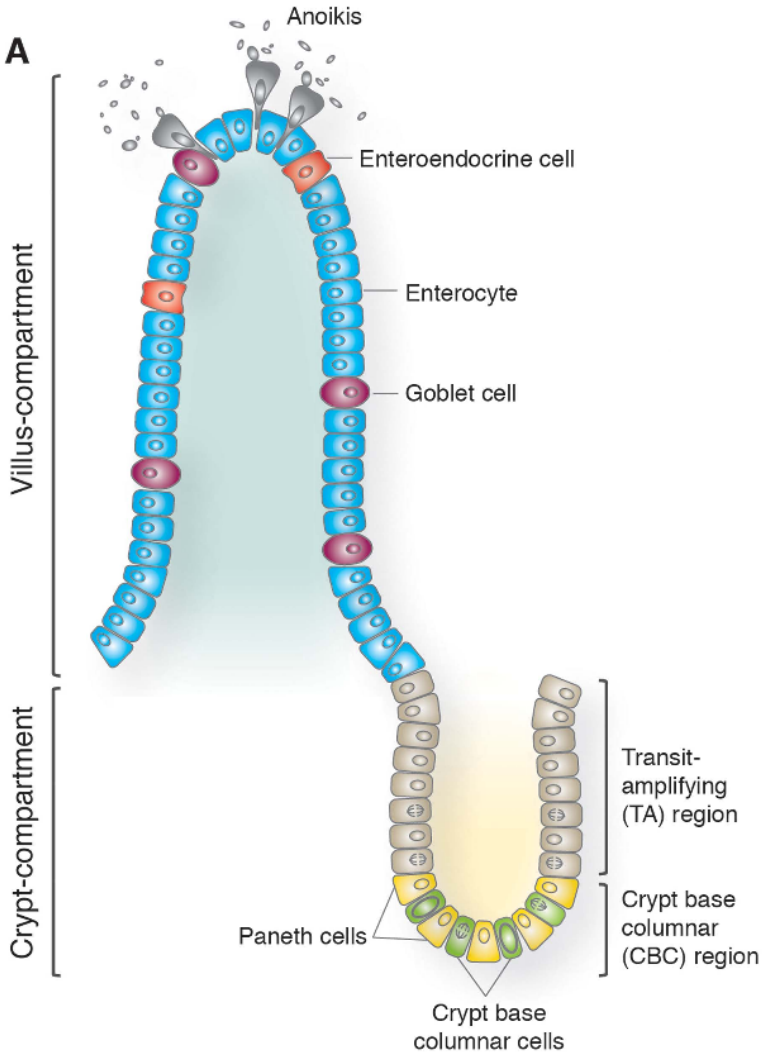


Figure 5: **The intestinal epithelium present a high turnover rate (~ 2-6 days) (adapted from Leushacke et al, 2014 [12])**. Progenitor cells are located at the bottom of the crypts and progressively migrate to the apex during maturation.

and glia that pattern concentrically within the wall of the bowel [16]. It has multiple roles: coordinating mobility across all GI tract, regulating gastric acid secretion and fluid movement, changing local blood flow and modifying nutrient handling [16–18]. Control can be **local** with specific enteric reflexes from ENS or **central** with the intervention of central nervous system (Figure 6). The importance of integration of signals between ENS and CNS differs along GI tract and the ENS organisation differs from conventional descriptions of the autonomic nervous system. Indeed, efferent pathways from CNS normally adopt two neurons in series structure [19] while the intestinofugal neurons from ENS are neurons with cell bodies in GI wall and axons projecting directly to sympathetic ganglia, pancreas, gallbladder and trachea but also to CNS [17] (Figure 6). In both small and large intestine, ENS uses full reflex mechanisms to direct movements in correct direction and to control fluid movements between the lumen and tissues [20]. Furthermore, ENS is in strong relationship with both immune and endocrine systems which leads to essential synergic roles in maintining mucosal barrier integrity and modulating nutrient absorption.

1.2.5 Enteric endocrine system

The imposing surface area of the GI tract consists mostly of absorptive enterocytes. Enteric endocrine system is represented by the large range of subsets of endocrine cells called Enteroendocrine cells (**EECs**). They represent about 1% of GI epithelial cell population and they are scattered along the tract together with other secretory cells [21]. All EECs collectively constitute the largest endocrine organ in the human body, in terms of number of cells and especially in terms of the variety of hormones secreted. They express more than 30 different hormone genes in all GI subsets combined [22]. They sense the luminal content and release corresponding hormones to act on the digestive processes locally or through direct connections with the nervous system to initiate appropriate responses in the global GI tract [19, 23]. The EECs originate from the same progenitors as other epithelial cells (Figure 7) and are also involved in the constant turnover from GI epithelium [21].

Depending on their localization, EECs hormone expression and morphology can differ to adapt to local GI functions. They are divided into two categories: “open” and “closed”

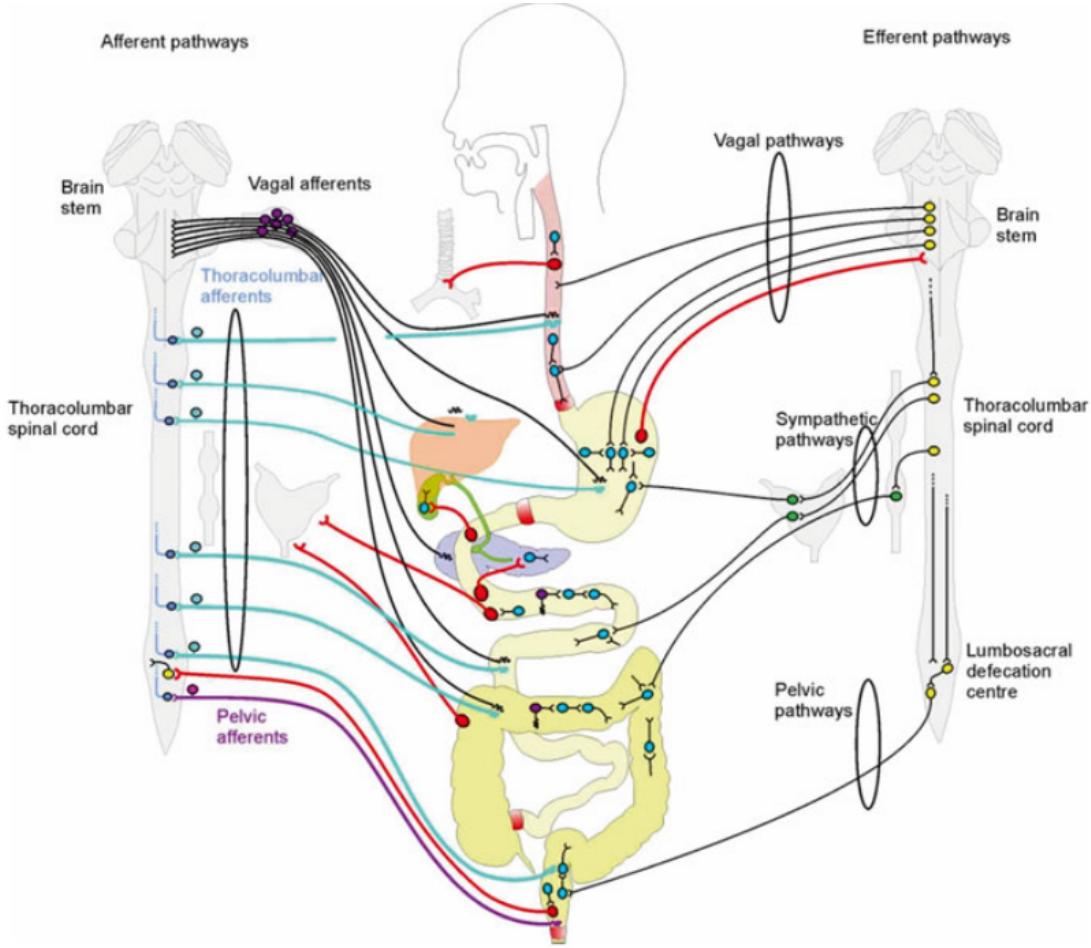


Figure 6: **Overview of the innervation of the gastrointestinal tract (from Furness JB, 2016)**. CNS and ENS work together or separately to control the digestive process tightly.

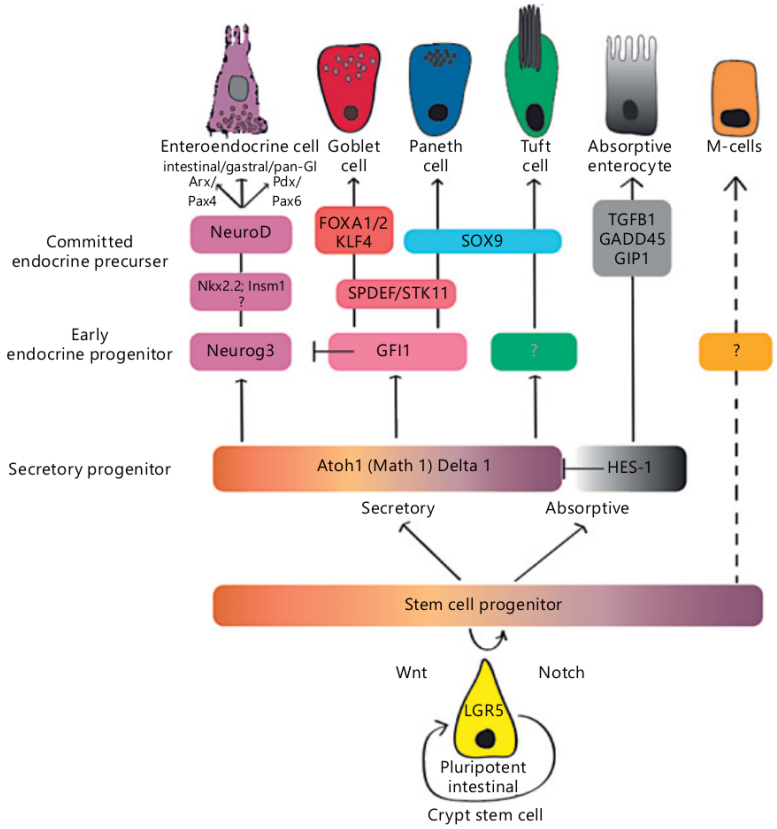


Figure 7: **Schematic overview of differentiation of intestinal epithelial cells in mammals (from Posovszky et al, 2017).** Early in the differentiation, secretory or absorptive fate are separated. Neurog3 expression initiates the enteroendocrine fate.

types. *Open-type* EECs present bottle neck shape and apical prolongation in order to reach lumen of the gut as they sport microvilli to directly sense luminal content. *Closed-type* EECs localize near the basal membrane and don't reach lumen which means that their activation is indirect [23, 24]. For example, *Gastrin* secretion is exclusive of *open-type* G-cells from stomach body and piloric region in response to luminal detection of aminoacids and calcium. *Gastrin* acts on *close-type* enterochromaffin-like cells (**ECL**) of gastric corpus which respond by producing histamine to activate secretion of acid gastric by parietal cells [25].

More than 15 different subtypes of EECs have been identified along GI tract [26, 27] (Figure 8). EECs are usually classified into subtypes according to the main hormone they secrete. This classification has been discussed and it has to be modified as extensive coexpression of hormones has been recently reported [28, 29]. Furthermore, all subtypes are not distributed equally in GI tract as the location and environment drive the regulatory mechanisms associated and required.

Secreted Product	Cell Type	Stimuli	Effect
Cholecystokinin (CCK)	I/CCK	Fat and protein	Increases gallbladder contraction and pancreatic secretion; decreases gastric emptying; anorectic
Gastrin	G	Food/nutrient/protein	Increases gastric acid secretion via ECL cells
Ghrelin	P/D1 (gastric), M	Fasting	Increases gastric emptying; orectic
Glucagon-like peptide 1 (GLP-1)	L	Fat	Decreases gastric emptying; incretin effect; anorectic
Glucagon-like peptide 2 (GLP-2)	L	Fat	Intestinal trophic factor; enhances digestive enzyme activity; decreases gastric emptying
Glucose-dependent insulinotropic polypeptide (GIP)	GIP	Fat	Incretin effect
Histamine	ECL	Gastrin	Increases acid secretion
Melatonin		Food/nutrient	Circadian entrainment; increases pancreatic secretion
Motilin	M	Fasting	Cyclic increase regulates migrating motor complex
Neurotensin	N	Fat	Increases pancreatic and biliary secretion and colonic motility; decreases gastric and small intestinal motility; anorectic
Oxyntomodulin (OXM)	L	Nutrients	Anorectic; incretin effect
Pancreatic polypeptide	PP	Food/nutrient	Decreases gastric emptying; anorectic
Peptide YY (PYY)	L	Fat	Decreases gastric emptying and small intestinal motility; anorectic
Secretin	S	Acid	Increases pancreatic secretion
Serotonin (5-HT)	Enterochromaffin, S	Luminal distention	Multiple effects on gastrointestinal motility
Somatostatin (SST)	D	Multiple, complex	Inhibits hormone and exocrine secretion
Uroguanylin		Salt	Regulates sodium homeostasis; mucosal protection

Figure 8: **Enteroendocrine Cell Products and Functions (from Shroyer et Kocoshis, 2011)**

1.3 Pancreas

1.3.1 Anatomical description and functional organization

Pancreas is one of the accessory organs of digestion along with tongue, salivary glands, liver and gallbladder. Located behind the stomach, it divides in 3 anatomical regions:

head of pancreas inserted inside the C-shape of duodenum; **body** representing the middle portion; **tail** of pancreas which extends to spleen region. It forms a glandular organ with amphicrine functions. Most of the pancreatic tissue is dedicated to pancreatic juice production[30]. The large network of ducts collects and drives juice into the duodenum lumen through duodenal papilla. Together they form the **exocrine pancreas**. On the other hand, **endocrine pancreas** represents a very small portion of the pancreatic tissue. However it produces and releases important hormones and performs fundamental role in blood glucose homeostasis.

1.3.2 Exocrine pancreas

Exocrine tissue represents around 95% of the whole pancreatic tissue and is in charge of pancreatic digestive contribution. It consists of two mains functional structures: **acini** and **ductal tree**. The exocrine fundamental unit is composed of one acinus and an intercalated duct. Acinar cells are organized in a lobule around an open chamber into which they release their secretions. Intercalated ducts converge from acini and merge progressively to form successively intralobular and interlobular ducts. It results into a principal duct (**Wirsung**) which meets the common bile duct to form the **ampulla of Vater** before flowing into the duodenum through *major duodenal papilla.* A secondary duct (**Santorini**) is sometimes observed. It has been observed that it can connect directly to the duodenum through *minor duodenal papilla* or drains into the principal duct. **Sphincter of Oddi** is the smooth muscular muscle forming a valve to control the flow of digestive juice from ampulla of Vater and it also avoids any reflux of duodenum content into the ductal tree [31].

Acinar cells are specialized for production and export of a large amount of proteins. They adopt a polarized morphology favorable to this function. The basal side of the cell contains large and developed endoplasmic reticulum which actively produces enzymes and pro-enzymes (**zymogens**) which are encapsulated inside vesicles and transported to the Golgi apparatus. Immature secretory granules leave the Golgi complex and the maturation occurs as they move closer to the apical side of the cell. Mature **zymogen granules**, marked by condensation of proteins, are more electron-dense and their size has decreased to two third [33]. Finally, **exocytosis** of secretory granules content is under the control

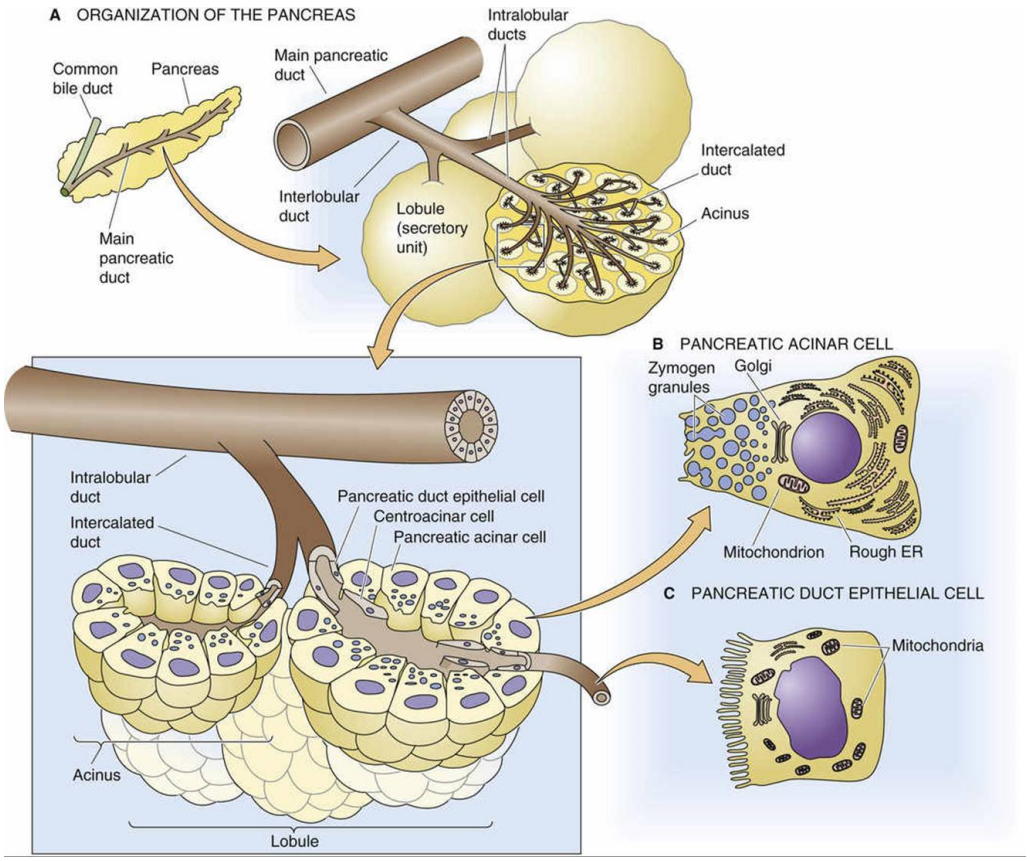


Figure 9: **Morphology of pancreatic acini and ductal tree (from Medical physiology - 3rd edition[32])**. Acinar cells are organized around an open chamber to form a lobule. Acini secretions are collected by the ductal tree which brings them to the duodenum. Both cell types present a polarized morphology in relation with their function.

of neurohumoral stimulation on the basal side of the cells [34, 35]. Cholecystokinin and gastrin produced by EECs have been associated with the hormonal stimulation of exocytosis. This induces the fusion of granules with the apical cell membrane followed by the release of digestive enzymes in acinus lumen. Pancreatic digestive enzymes include **trypsin**, **chymotrypsin**, **pancreatic lipase**, **amylase** but also ribonuclease, deoxyribonuclease, gelatinase and elastase [36]. Trypsin and chymotrypsin are proteases synthesized in their inactive form, **trypsinogen** and **chymotrypsinogen** respectively, in order to avoid protease activity inside acinar cells. Trypsinogen will be activated into trypsin by *enterokinase* expressed at the surface of the intestinal epithelium. Trypsin activates successively the conversion of chymotrypsinogen into chymotrypsin together with additional trypsinogen molecules.

Centroacinar cells ensure the junction between acini and the ductal tree and are followed by ductal cells which adopt more classical cuboïdal morphology [31]. Pancreatic duct cells produce an ion-rich fluid, increasing the pancreatic juice volume, which hydrates and alkalinizes protein-rich secretions of acini through secretion of water and and bicarbonate ($NaHCO_3$) [37–39]. The EECs inside duodenum produce a hormone, secretin, which stimulates the production of bicarbonate by the ductal cells [40]. When flowing in the duodenum, pancreatic juice neutralizes gastric-inherited acidity of the chyme in order to allow pancreatic enzymes to work optimally for the digestion of nutrients. Finally, pancreatic ductal cells also secrete glycoproteins in order to protect the ductal tree from enzymatic activity.

1.3.3 Endocrine pancreas

The endocrine tissue represents 1-2% of the whole pancreatic tissue [41]. All pancreatic endocrine cells (**PECs**) are regrouped inside encapsulated small structures, called **islets of Langerhans**, which are scattered between exocrine tissue with a higher density detected in the tail of pancreas in human. Endocrine cells represent the great majority of islet cells and numerous other cell types are represented: endothelial cells, other stromal cells, immune cells, and neural elements [42, 43]. Vascularization of islets is strongly developed according to the role performed by endocrine cells in sensing blood content. PECs release the hormones directly into the blood, regulating distant target

organs. While they represent only 1-2% of pancreatic mass, islets receive 10 to 15% of blood flow dedicated to whole pancreas [43]. This is achieved through the formation of a glomerular-like structure of fenestrated capillaries which allows close relationship with endocrine cells.

We distinguish five major PECs depending of the main hormone they produce: **beta cell** (β) producing *insulin*, **alpha cell** (α) and *glucagon*, **delta cell** (δ) and *somatostatin*, **gamma** (γ or **PP**) **cell** producing *pancreatic polypeptide* and finally the *ghrelin*-expressing **epsilon cell** (ϵ) [44] (Figure 10).

In human, alpha (~40%) and beta (~50%) cells represent around 90% of PECs. A large interspecies comparison of the islet structure has demonstrated that the proportion of alpha and beta cells varies between species (Figure 11) [45]. Moreover, variations have also been observed in the organization of the different endocrine cells within islets.

Insulin and *glucagon* are the two key regulators of the glucose homeostasis and they will be discussed in the next section. *Somatostatin* (*sst*) is a widely effective hormone in human body which acts as general inhibitory hormone in several organs. In the digestive tract, it has global inhibitory effect on several processes: *Sst* is an inhibitor of the gastrin release and thus the gastric acid secretion; it inhibits both bile and pancreatic secretions; it has been described that *sst* inhibits the intestinal nutrient absorption but also the intestinal mobility; finally it acts as an inhibitor of both *insulin* and *glucagon* secretions [46–49]. Pancreatic polypeptide (PP) acts as a postprandial important feedback hormone and it has been associated with the control of appetite, weight and both endocrine and exocrine pancreatic secretions [50–52]. Finally, *ghrelin* is known to be the hunger hormone but recent findings highlighted *ghrelin* roles in glucose homeostasis, energy homeostasis, cardiac functions, muscle atrophy and bone metabolism. The presence of *ghrelin* hormone and its receptor are also detected in many tumors where they might be associated to the tumor growth and progression [53].

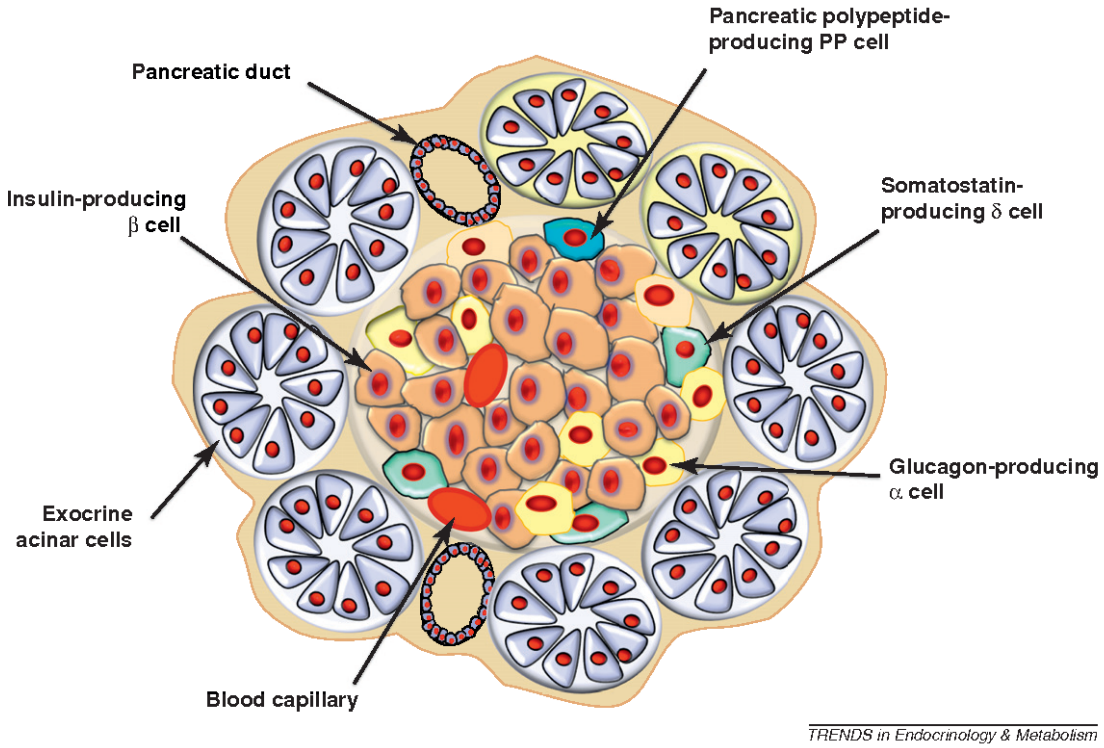


Figure 10: **The organization of the endocrine tissue**. Islets are clusters of endocrine cells and they are dissiminated inside the pancreatic exocrine tissue. Islets are vastly vascularized as endocrine cells act as sensor of blood content.

1.4 Glucose homeostasis

Glycemia is the glucose concentration in blood circulation. Glucose homeostasis consists of all mechanisms activated dynamically to maintain this concentration in physiological range (80-100 mg/dl) [54]. As multiple events affect glycemia, this means that the regulation is constant and highly dynamic. The balance between *insulin* and *glucagon* secretion is the crucial driving mechanism for blood glucose homeostasis. *Insulin* is the hypoglycemic hormone secreted by pancreatic β cells in response to an increased glycemia. β cells present glucose sensing ability through expression of the glucose transporter SLC2A2, glucokinase and through the glycolytic and oxidative metabolism of this sugar [55, 56]. ATP-sensitive K+ (KATP) channels play central role by linking the insulin secretion to the increase of blood glucose (glucose-dependent insulin secretion). At physiological glucose concentrations, KATP channels are open and K+ efflux creates a negative potential on the cell membrane which keeps the voltage-gated Ca2+ channels closed. A higher level of glycemia increases the glucose uptake by beta cells. It stimulates

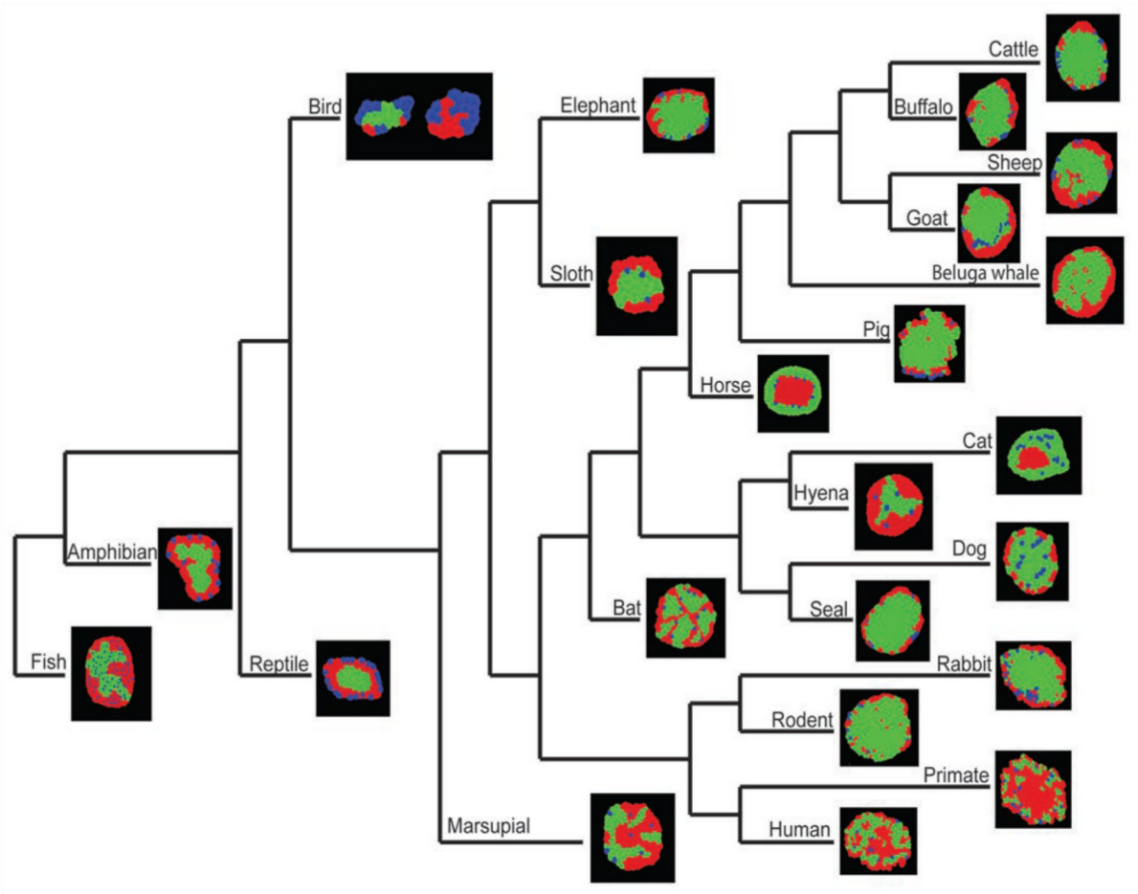


Figure 11: **Multi-species comparison of the islet structure and composition (from Steiner et al, 2010)**. The proportion of each endocrine cell type is subject to changes between the different species. Alpha cells (red), Beta cells (green) and Delta cells (blue).

the metabolism of beta cells resulting in changes in the cytosolic content, such as an increase of ATP cytoplasmic concentration, and leading to the closure of KATP channels [57]. The resulting depolarization of the membrane induces the opening of Ca2+ channels and leads to a Ca2+ influx into the cell. The rise of Ca2+ cytoplasmic concentration triggers exocytosis of secretory granules containing insulin [56, 58]. Recent evidence has shown that a high glucose level in beta cells also induces the internalization of KATP channels, and this could play an even greater effect on the excitability of beta cells than the ATP-dependent closure of KATP channels [59]. Conversely, glucagon is an hyperglycemic hormone. Alpha cells present similar mechanisms of release at the difference they react to a decreased glycemia [60].

Glucose is the fundamental energy unit for cell activities and the cell glucose uptake occurs through two mechanisms, called insulin-mediated glucose uptake (IMGU) and non-insulin-mediated glucose uptake (NIMGU) [61]. *Insulin* affects preferentially skeletal muscle, liver, adipose tissue and nervous system by stimulating the recruitment of glucose transporters from cytoplasmic region to the plama membrane of cells [62, 63]. The glucose excess is captured by these tissues and is successively metabolized for glucose storage: **glycogenesis** (glycogen synthesis) in myocytes and hepatocytes; **lipogenesis** in adipocytes and hepatocytes. Conversely, *glucagon* is expressed during physical activities or between two meals, events characterized by decreasing glycemia. *Glucagon* triggers **glycogenolysis** in liver which is the catabolism of glycogen molecules into glucose units. The release of these units into blood circulation compensate the decrease in the glucose concentration (Figure 12).

Furthermore, *Insulin* and *glucagon* act directly on opposite cell types to inhibit the antagonist secretion. This balance is also controlled by other hormones or signals from body which implies that any disturbance can have major repercussions.

Finally, several different studies have identified a group of genes whose expression is systematically reduced in pancreatic islet cells [64–68]. The expression of these "disallowed" genes in pancreatic islet cells alters some of their important functional properties, such as the normal regulation of insulin secretion in beta cells. The importance of the repression of these genes in pancreatic islet cells is emphasized by the observation that a disregulation of their expression is often observed in type 2 diabetes [69]. Moreover, the

emergence of RNA-Seq datasets from purified islet cells has allowed to study the specific repression of these disallowed genes in each individual islet cell types [70]. As a result, the most strongly disallowed genes are similar in the different islet cell types. However, several genes are only repressed in beta cells or alpha cells and could be associated with more specific functional properties of these cells.

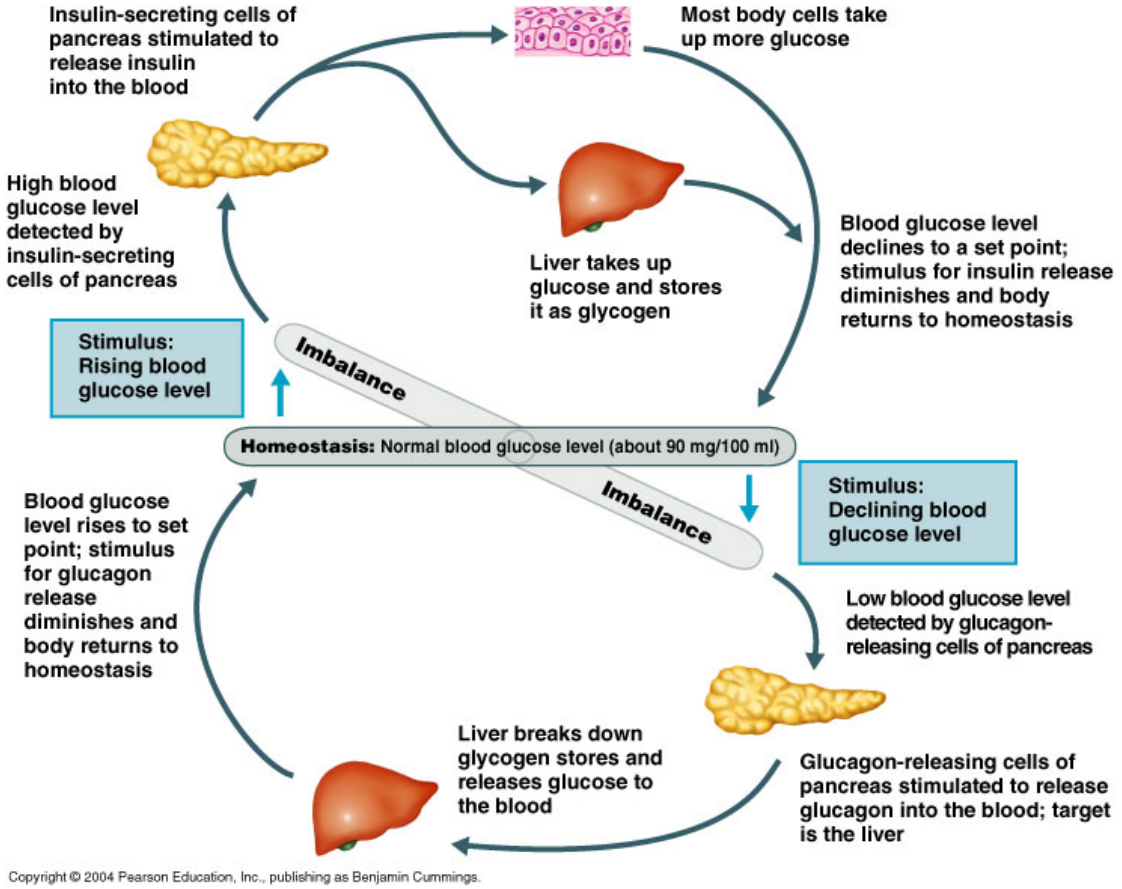


Figure 12: **Glucose homeostasis**. Imbalance in glycemia will trigger the response of the corresponding cell type to bring back normal blood sugar. Insulin increases absorption of glucose by the liver and the rest of the body tissues to decrease the glycemia. Conversely, glucagon can only stimulate glucose release from the liver to increase glycemia.

1.5 Diabetes

Diabetes is the term describing heterogeneous metabolic diseases where patient presents hyperglycemia over prolonged period. Diabetes is often reduced to **Diabetes Mellitus** (**DM**) in which the glucose homeostasis integrity is altered. However, it exists a second form, called **Diabetes insipidus**, which originates from neurogenic or nephrogenic defects of the vasopressin (ADH) signalling and resulting in excessive water loss by kidney [71]. **DM** is divided in different categories based on the origin of the hyperglycemia. Type 1 diabetes (**T1D**) covers all cases where insulin secretion is altered or missing. It originates from an auto-immune disease which depends on genetic predispositions or/and environmental conditions. Beta cells destruction is observed in T1D patients and lead to a complete beta cell mass loss. Type 2 diabetes (**T2D**) is generally associated with tissue insulin resistance disrupting **IMGU**. Evidence has shown that most of T2D originate from two major pathophysiologic abnormalities [72]. (1) The gradual increase of the peripheral resistance to insulin which can be due to defective insulin signaling pathway or high lipidemia. In early stages of the developping disease, beta cells increase insulin production to overcome the insulin-resistance and some patients will never develop symptoms, even if no adaptation of diet or physical activities occurs. When beta cells overload becomes unsustainable, beta cells destruction is observed together with development of symptomes. (2) The beta cell failure plays a key role and can be observed in the early stages of T2D development, already leading to a reduced expression of insulin while a significant peripheral resistance is yet to be observed [73]. It has been associated with a reduced glucose sensitivity of the beta cells. Their functional activity could continuously decline in parallel with the developement and the progression of the disease. Genetic component is observed while an obesity associated with a fatty-rich diet and a lack of sportive activities are two strong risk factors for developping T2D. T1D and T2D accounts for ~10% and ~90%, respectively, of diagnosed diabetes [74, 75]. Beside those two, other forms of diabetes can be observed: Gestational Diabetes (GD) can be developed during pregnancy; Maturity Onset Diabetes of the Young (MODY) covers inherited autosomal dominant mutations of genes which directly affect insulin production/secretion.

2 The Zebrafish model

The use of animal models in biological and medical research is a common approach. Rodents, such as mouse (*Mus musculus*) and rat (*Rattus norvegicus*), are widely used as classical vertebrate models. Many strains have been actively selected and developed due to their ability to mimic human traits, often associated to particular diseases. During the last 3 decades, the use of the zebrafish model (*Danio rerio*) has emerged quickly following major breakthroughs in genetic engineering methods [76]. Zebrafish is a small-bodied tropical, freshwater fish species from South Asia with several key features such as the ease of care and housing, year-round prolific breeding and its large offspring which make it an interesting affordable model for laboratories. Beside the practical benefits, its biological features contributed widely to its recent success in many fields of biology. External embryonic development, quick organogenesis, associated with the transparency of the embryo and surrounding chorion make it an excellent model to study development, notably using fluorescent transgenic lines (Figure 13). As previously mentioned, progress in molecular biology, high-throughput sequencing and imaging techniques increased the efficiency ot tools used on zebrafish and prompt many scientists to work on that animal model. Indeed, the relative ease of genetic manipulation in zebrafish has led to a wave of gene function studies by genetic screening of mutants at large scale. This infatuation for zebrafish was the so called “Big Screen” initiative undertaken by Christiane Nusslein-Volhard in 90s [76, 77]. This effort has been followed by the sequencing of the zebrafish genome between 2001 and 2013 which highlighted about 26000 protein-coding genes in zebrafish [78]. In recent years, forward genetics has been progressively replaced by targeted mutagenesis notably due to the emergence of new methods such as TALEN [79] and CRISPR/Cas9 [80] which contributed to make zebrafish an interesting tool for disease modeling.

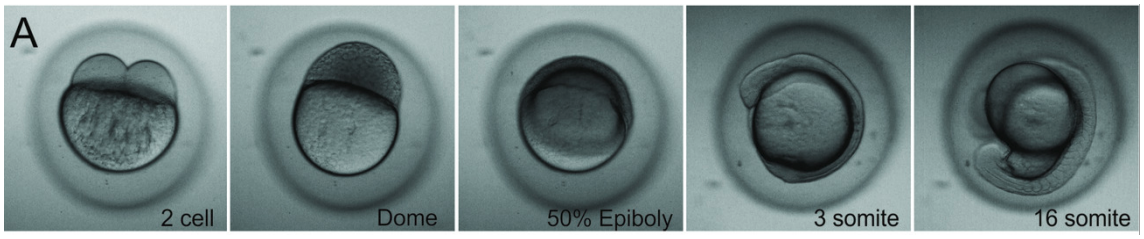


Figure 13: **Zebrafish external development** (from Meyers et al, 2018). During first stages of its developement, the zebrafish embryo is translucent. It allows to observe and study the development.

2.1 Comparative genomics of zebrafish

Zebrafish and human divergence time estimation is 450 mya (million years ago) when bony vertebrates (*Osteichthyes*) divided into **ray-finned fishes** (*Actinopterygii*) and **lobe-finned** fishes (*Sarcopterygii*). Zebrafish belongs to **Cyprinidae** family inside **Teleostei** infraclass, the largest of ray-finned fishes. The teleosts (~ 310 mya [81]) experienced an additional whole-genome duplication (WGD) to the two rounds of WGD experienced by all vertebrates, with exception of the superclass Agnatha [82]. Comparative genomics among the teleosts suggests that this teleost genome duplication (TGD) occured at the base of the teleost radiation [83]. However, it is believed that whole genome chromosomal rearrangement occurred in the long lineage of teleosts leading to zebrafish. This hypothesis is supported by the fact that the zebrafish genome only retained 20 to 30% of the duplicated genes pairs from TGD. Moreover, zebrafish genome presents a similar amount of chromosomes as human (25 and 23 pairs respectively), rather than twice as many, as would be expected from the TGF.

The identification of conserved syntenies between zebrafish and human genomes was held back by genome rearrangement and chromosomal restructuring between both species [84]. The recent sequencing of the genome of **spotted gar** (*Lepisosteus oculatus*) [85], a ray-finned fish diverging from teleosts prior to the TGD, revealed higher conservation between human and gar compared to conservations between zebrafish and human or zebrafish and gar. This observation supports the hypothesis that the TGD accelerated the loss of ancestral syntenies in teleosts through chromosome rearrangement. WGD is believed to be one of the major evolutionary events facilitating syntenic rearrangement [86].

Conserved non-genic elements (CNEs) are invaluable resources in comparative genomics as they represent non-coding regions selectively conserved. They have been associated with transcription factor binding sites and other cis-acting regulatory elements and they are potential important sites of evolutionary divergence. The gar genome also bridged the teleosts to the **tetrapods** (including human) by highlighting novel syntenic regions, including CNEs, between teleosts and human that do not directly align between them [85]. The identification of CNEs by pairwise comparison of largely evolutionarily distant species is challenging whereas comparing multiple genomes is more robust at capturing orthologous sequences.

Finally, zebrafish is evolutionarily distant from any other available fish genomes leading to its current phylogenetic isolation [87] (Figure 14-A). The distance between zebrafish genome and the closest sequenced teleost is even higher than the distance between human and chicken genomes (Figure 14-B). By identificating and sequencing a fish at a comparable molecular distance to that between human and mouse, additional CNEs currently not annotated in zebrafish genome are expected to be revealed in the future.

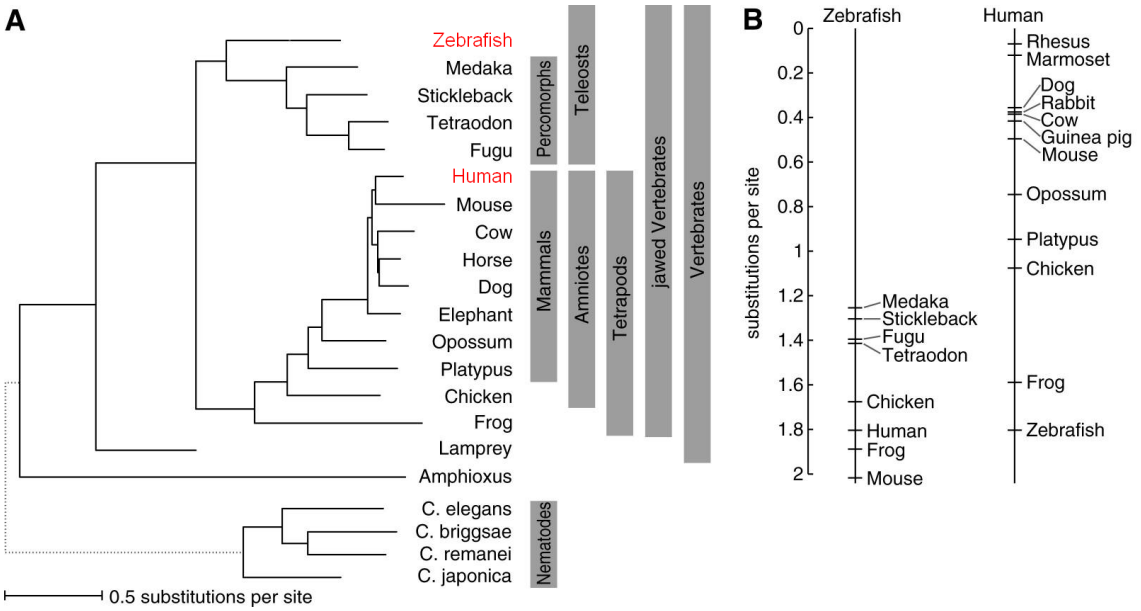


Figure 14: **The evolutionary distance between zebrafish and human (adapted from Hiller et al, 2013 [87])**. (A) Phylogenic tree with branch lengths. Zebrafish is currently classified as a phylogenetic isolated species. (B) Evolutionary distances (neutral substitutions per site) of sequenced species compared to zebrafish and human respectively. This highlights the absence of sequenced organisms at a close molecular distance from zebrafish.

2.2 Zebrafish and Gastroinstestinal tract

The zebrafish owns a simplified version of the mammals GI tract highlighted by the lack of a proper stomach. Intestine fills majority of the abdominal cavity and is divided in three regions: the anterior, the mid and the posterior intestinal segments (Figure 15). The anterior intestinal segment, often called "the intestinal bulb", presents wider caliber than any other portions of the intestine and thus may function as a reservoir [88]. Similarly to mammals, GI functions are spatially distributed along the tract. Anterior and mid intestinal segments are responsible on protein and lipid absorption, respectively, while ion transport and water absorption occur in posterior segment [88, 89]. Zebrafish intestinal epithelium contains the same cell types as found in mammals, except that it does not contain Paneth cells. Also, while the anatomy and architecture of the zebrafish gut are very similar to the mammalian gut, there are some differences such as the absence of the crypts of Lerberkuhn [88].

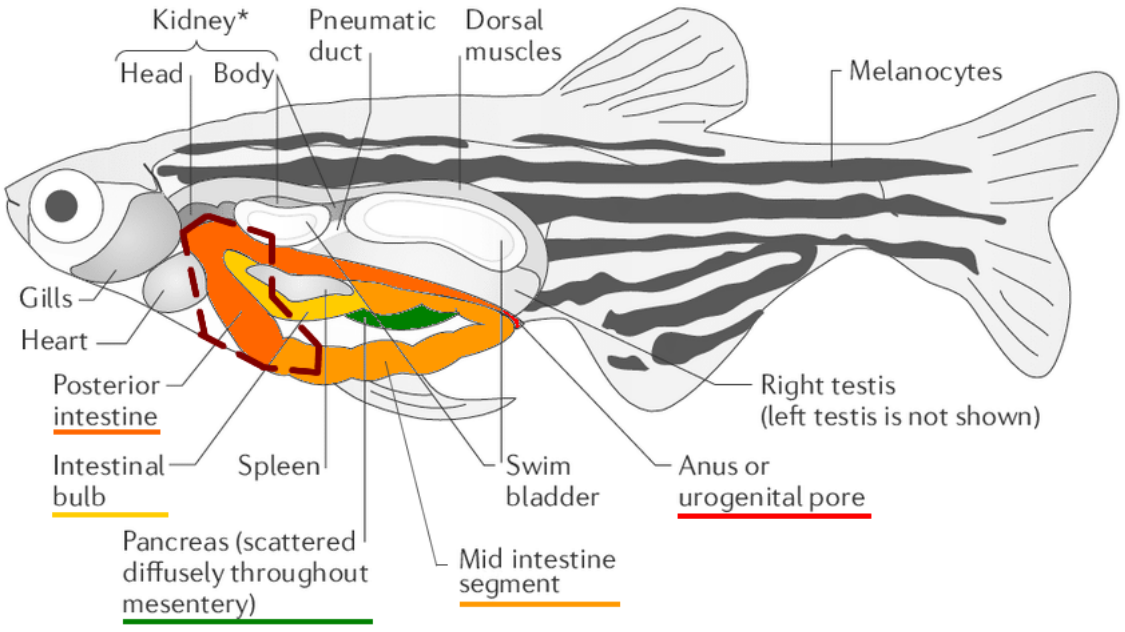


Figure 15: **The anatomy of the zebrafish GI tract (adapted from White et al, 2013 [90])**. The three intestinal segments are represented along with the pancreas. The red dotted area represents the liver region. GI tract ends with the anus located anteriorly to anal fin.

2.3 Zebrafish and Pancreas

The pancreas of zebrafish presents anatomical and functional organization similar to the mammalian pancreas. It is composed of pancreatic endocrine islets distributed inside exocrine tissue, composed by acini and ducts. However, the pancreas does not assemble in a concrete and compact individual organ but instead it forms a relatively diffuse tissue structure closely associated to the intestine (Figure 15). Exocrine tissue organization is strongly similar to mammalian structure as ductal tree consists of extra and intrapancreatic ducts connecting and collecting the secretions from acini [91]. However, the endocrine tissue is slightly different. Indeed, zebrafish endocrine pancreas consists of a big principal islet located in the pancreatic head as well as smaller secondary islets scattered in the tail segment [92]. Histologically, islets contain same cell subtypes with the exception of PP cells which not exist in zebrafish.

Pancreas functions are comparable to those found in mammals as evidenced by the presence of active regulation of blood glucose through the balance between glucagon and insulin secretions [93].

3 Next-Generation Sequencing

Next generation sequencing (NGS) or high-throughput sequencing (HTS) refers to all technologies that perform massively parallel sequencing, producing millions of sequences concurrently [94]. It has revolutionized genomic research by allowing the sequencing of genomes within a single day and by breaking "the $1000 genome" challenge [95]. Last 10 to 15 years, NGS technologies have contributed to decrease drastically the sequencing costs. They have driven the quick emergence of a multitude of applications with the most recent advances allowing researchers to work at the resolution of a single cell (Figure 16 and see https://www.illumina.com/content/dam/illumina-marketing/documents/applications/ngs-library-prep/ForAllYouSeqMethods.pdf). In this section, some popular applications will be presented as they were used in this work.

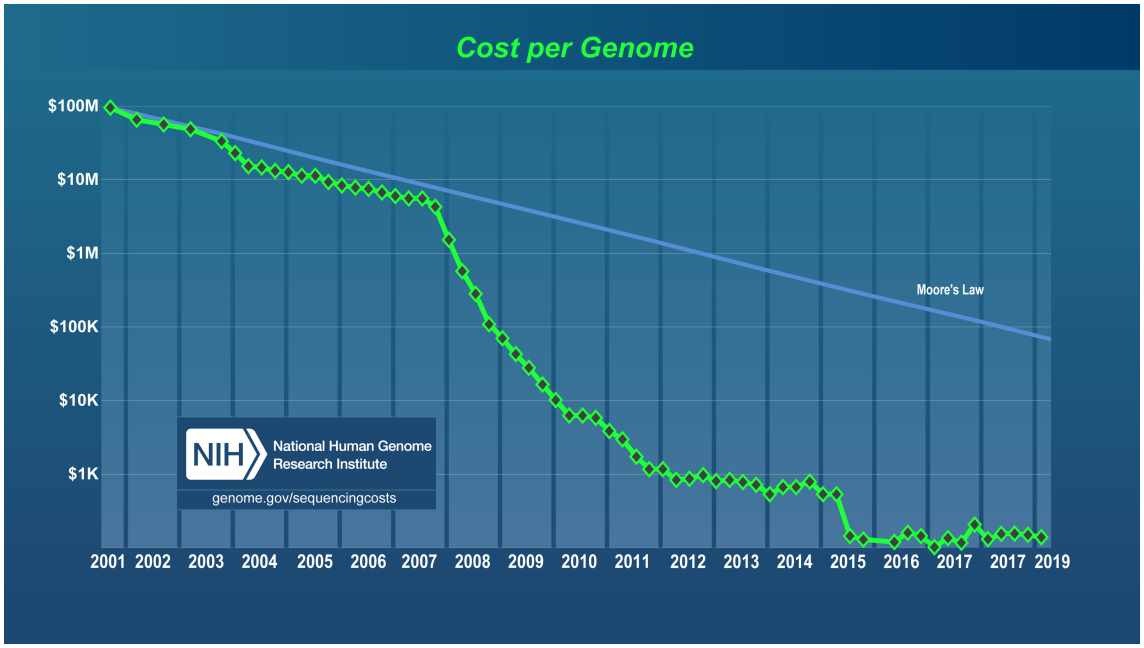


Figure 16: **Genome sequencing costs** (*https://en.wikipedia.org/wiki/$1,000_genome*). In the past two decades, the evolution of technologies has decreased greatly the sequencing cost per genome.

RNA Sequencing (RNA-Seq)

RNA-Seq is one of the most popular derived technology from NGS/HTS. The most common usage is to study the cell transcriptome by quantifying RNA presence in biological samples. RNA can be isolated from heterogenous tissues or from a specific cell type using cell sorting. RNA is purified to remove any DNA content. Quality of isolated RNA is

then monitored as it will influence the following steps. Depending on the interest, the selection or depletion of specific type of RNA can be performed: messenger RNA selection, ribosomal RNA depletion, RNA capture or size selection for micro-RNA. Selected RNA is then reverse transcribed into more stable cDNA. Fragmentation of cDNA is performed and is followed by the size selection of fragments with convenient size for the chosen sequencing technology. Library preparation is the final step before the sequencing as specific adaptators will be added on both sides of fragments to allow the sequencing process.

Chromatin Immuno-precipitation Sequencing (ChIP-Seq)

This method combines chromatin immunoprecipation and HTS. It is mainly used to study the interaction of a protein with specific DNA regions. Using freshly dissected biological materials, the crosslinking of protein-DNA complexes is performed using formaldehyde solution. Complexes are then sheared into DNA fragments (~ 200-500 bp) by sonication or nuclease digestion. Immunoprecipitation is performed using specific antibodies targeting complexes formed by the protein of interest. Crosslinking is reversed by heating and DNA molecules are purified and libraries are prepared for sequencing.

Single Cell Sequencing

Most recent optimizations of NGS capabilities allow now to define signal at the level of a single cell. RNA Sequencing at the level of a single cell (scRNA-Seq) has been the first application but protocols for ATAC-Seq or Whole-genome-amplification have also been adapted to single cell resolution. scRNA-Seq allows highlighting and describing the transcriptome of each existing cell type inside heterogenous tissue samples. It presents the advantage to avoid the cell sorting step which has been the major concern to study transcriptome of heterogeneous cell populations. However, the technique is less sensitive than bulk RNA-Seq with the current estimation that around 15% to 20% of RNA content is captured for a single cell. Finally, working with thousands of unique cell transcriptomes in parallel has considerably increased the needed computational power and bioinformatic software applied to single cell technology have developed quickly in recent years.

4 Data analysis of transcriptomic experiments

4.1 The importance of the experimental design

The success of all RNA sequencing technologies is due to the improved sensibility and the better ability at detecting transcripts than the previous methods, such as microarrays. It is mainly due to a higher sensitivity and it detects more dynamically the gene expression [96]. This sensitivity means that a special attention is required for the planification of the global experiment: identifying the correct control condition, defining the required amount of replicates, avoiding any confounding factors and listing all important parameters to control during the experiment. It is important that the investigators clearly define the appropriate experimental design for their research. In a well-designed experiment, the investigators should be aware of the characteristics and the weaknesses of their dataset to correctly adjust their conclusions and findings.

However, the increased sensitivity also means that all the small and uncontrolled changes during the experiment could affect the outcomes of the analysis. The interpretation of findings depends on the ability to minimize all the external sources of variability. The variability unrelated to the experimental design of the study is commonly called "batch effect". Koch and collaborators reviewed the most common sources of batch effect and they proposed strategies to minimize them [96]. However, it is not always possible to remove all sources of batch effect during the experiment. Using an appropriate software, it is possible to control their impact during the in-silico analysis. This can only be performed if the sources of variability have been clearly identified.

4.2 Aligning the data to a reference

Nowadays, most of the NGS facilities send the data to the researchers under FastQ format [97]. It contains the sequence of all the fragments "read" by the HTS platform for a given dataset, with the quality scores. When available, the reference genome is used to align all these sequences on their source location, and they are mapped to the set of known genes to obtain a raw count table. Specific pipelines allow the direct identification of transcripts and to predict alternative splicing.

The raw count table is the major output from this step. One could also use SAM/BAM files [98] to directly visualize the alignment using a genome browser, such as IGV [99–101].

4.3 Normalization, sample variability, and outliers

The "downstream analysis" usually refers to all the analysis performed after getting the raw count table. Based on the experimental design of the experiment, a variety of different analysis can be performed. However, some key steps are mandatory to optimally extract the information from a dataset.

The normalization of the different samples within a dataset is particularly important. The simplest approach for the inter-sample normalization is to scale all gene expression levels based on the difference of library sizes between samples [102]. Depending on the purpose, one could use different advanced methods to normalize the raw counts between samples [102]. Some factors such as the gene length or the global distribution of gene expression are then considered in non-linear normalization strategies. One could mention the "Reads Per Kilobase per Millions mapped reads" (RPKM) approach which is a quite popular method to normalize raw counts for sequencing depth and gene length. This is an efficient intra-sample normalization procedure. However, this method has been characterized to introduce a bias in the per-gene variances, in particular for lowly expressed genes, when performing a differential expression analysis [103].

The next important step is to assess the variability inside the dataset. Different methods can be used to assess this variability. The most used approaches are the Principal Component Analysis (PCA) and the different measures of correlation such as the Pearson's coefficient or the Spearman's rank correlation coefficient [104, 105]. PCA is a method from multivariate statistics for the dimensional reduction of data without the loss of information. It reduces all the variables of a dataset, usually all the genes, to a limited number of dimensions called Principal Components (PCs) still reflecting the complete variability inside the dataset. The PCs are ranked based on the variation they explain. PC1 explains the most variation within the dataset, PC2 the second most, and so forth. Both the PC1 and the PC2 usually cover together the majority of the total variation. The two-dimensional visualization of PCs is generally called PCA plot (Figure 17). The

measure of correlation between samples is another approach commonly used and it relies on the calculation of the distance between samples. This distance is represented by a correlation coefficient depending on the correlation method used. The higher is the coefficient between two samples, the most similar they are.

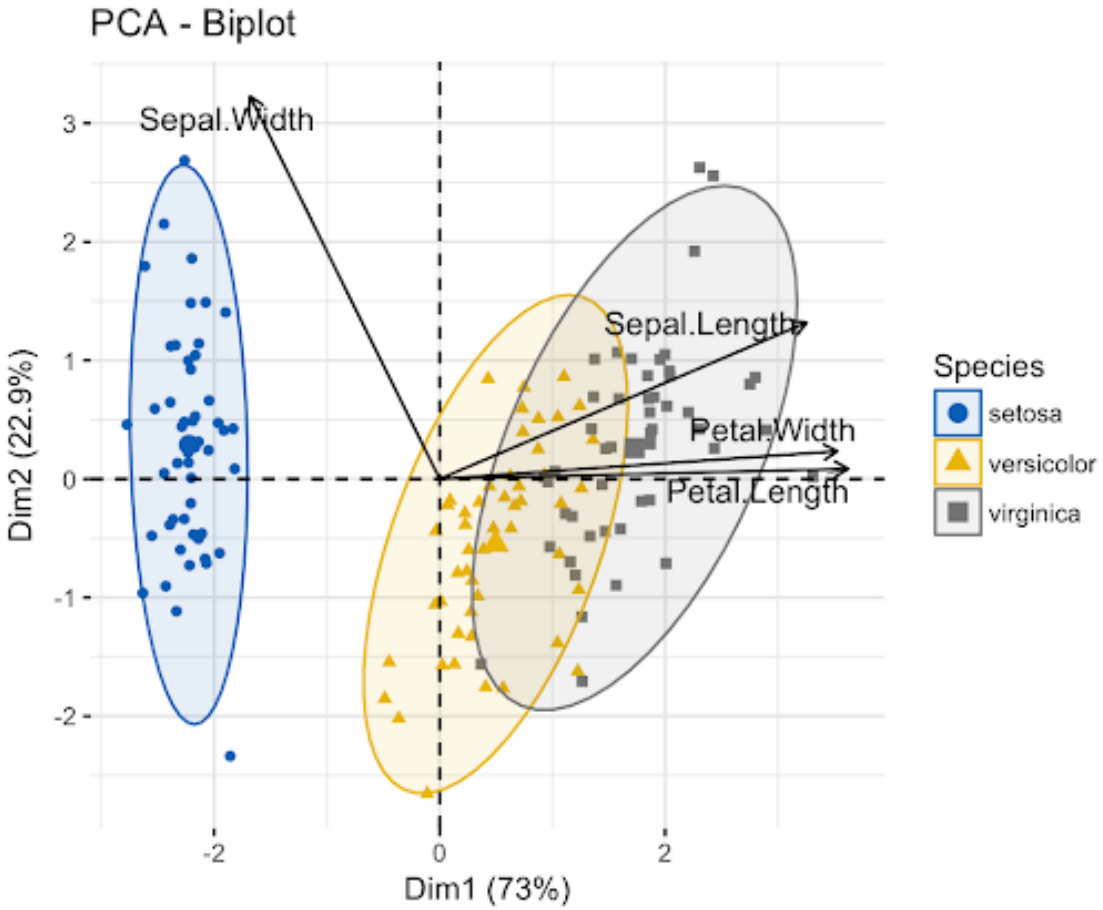


Figure 17: **PCA plot**. This is an example of the two-dimensional visualization of principal component analysis on the Iris Data Set. The contribution of original variables to each PCs can be observed.

Both approaches are tools to visualize the clustering between the conditions and the replicates to directly identify the potential outliers. The variability between the different conditions, the intergroup variability, is expected to be higher than the variability between samples inside a condition, the intragroup variability. In a well-design experiment, the intragroup variability is correlated to the residual batch effect. It may originate from both technical and biological variabilities. Moreover, the different experimental conditions may present a different intragroup variability as they can be differently affected by the external sources of variability. At the end of this process, one should be aware of the

variability inside a dataset and any clearly identified outliers must be removed to secure the downstream analysis.

4.4 Differential expression analysis

Once the quality control, normalization, and filtering steps are performed, the investigators may focus their analysis on the principal purpose of RNA-Seq analysis: the identification of a set of differentially expressed (DE) genes for the different conditions. Several tools have been developed in the last decade, presenting their own methods. Basically, they are quite similar and the choice of one or another may rely on the experimental design of the experiment. For example, the methods provided by the R packages DESeq2 and edgeR are very popular. Both rely on the strong assumption that most of the genes are not differentially expressed, and that for those differentially expressed there is an approximately balanced proportion of over- and under-expression [102, 106–108]. However, in most of these methods, the differential expression analysis is performed following a pairwise comparison between two conditions. When comparing multiple conditions (n > 2), one may choose an alternative approach using an analysis of variance (ANOVA) method.

The output of the differential expression analysis is a table. For each gene, it contains the log2 fold change between the two conditions (e.g. pairwise), the P value of the test, and the adjusted P value for multiple testing. The volcano plot is a good tool to visualize the table by plotting p-values against log2 fold change. However, the MA plot is an alternative vastly used to visualize DE genes (Figure 18). It consists in plotting log2 fold change against the mean of expression level over the two conditions (e.g. normalized counts, logarithmically transformed of not).

4.5 Other downstream analysis

From a list of DE genes, several additional downstream analyses may be performed depending on the interest of the investigators. One could be interested in clustering the genes following their pattern of expression. Hierarchical and k-means clustering methods are among the most frequently used. The purpose of these analyses is to identify different

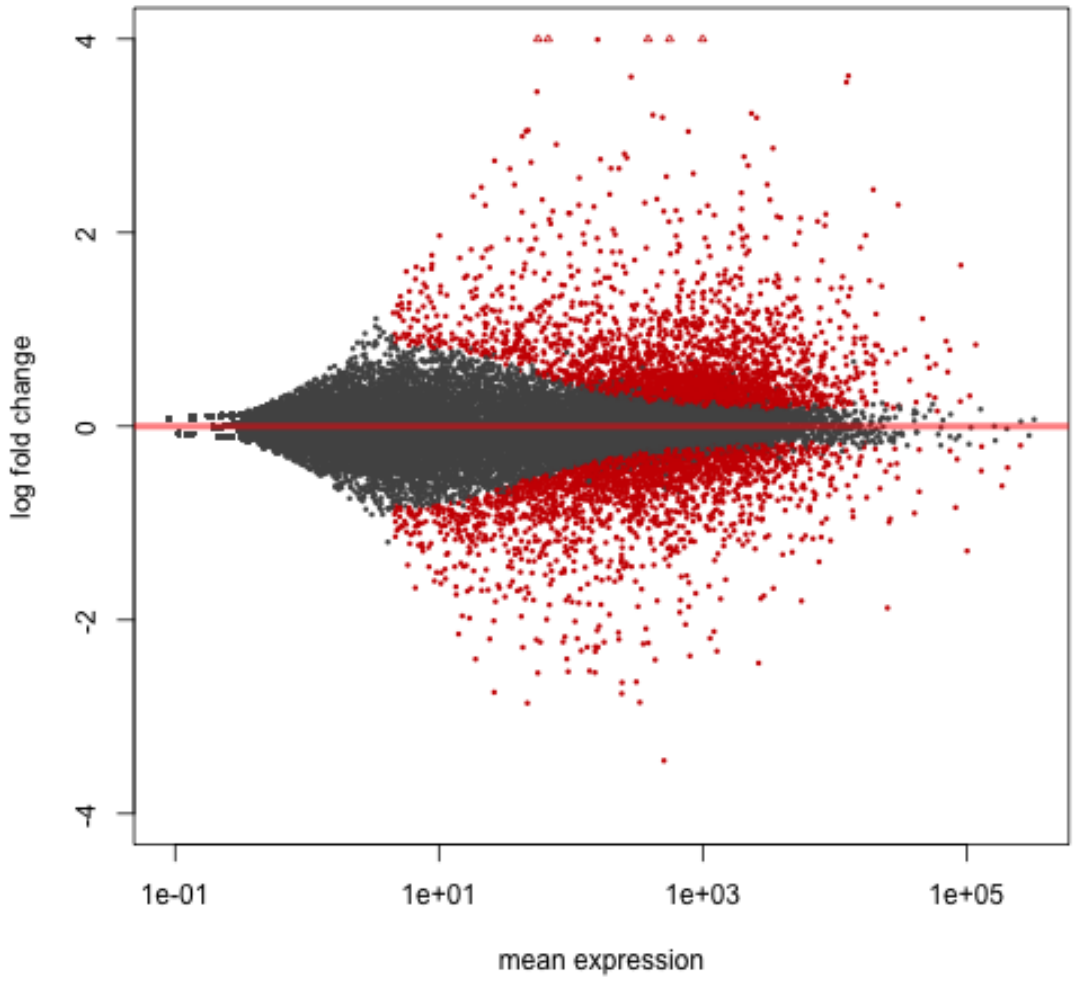


Figure 18: **MA plot**. This is an example of MA plot for the visualization of a differential expression analysis. For all genes, the log of the fold change between the two evaluated conditions is plotted in relation to the expression level. Red dots are the genes with a significant differential expression. Genes up-regulated in the treated condition are above the red line, genes down-regulated are below this line.

modules of genes with a similar expression pattern across the samples and conditions. Functional enrichment analysis represents the different methods for imputing a biological meaning to a set of DE genes. The approach relies on testing a list of genes against databases of biological information. Multiple "command-line" or online software (PANTHER [109], GOrilla [110, 111], GSEA [112, 113] or Enrichr [114, 115]) allow to use a variety of scientific databases such as Gene Ontology (GO), WikiPathways, KEGG or Reactome. Most of these methods perform a hypergeometric test to evaluate the significance of the enrichment.

4.6 Working with single-cell data

Working at the level of a single cell transcriptome brings an additional challenge. Prior to the downstream analysis, it already implies few changes to produce the gene count matrix. One of the most significant change is that you must attribute all reads to the all individual cells. This is achieved through the combination of cell barcodes and Unique Molecular Identifier (UMI). The approach is quite similar between the different single-cell technologies, with some specific methodological differences. UMIs are used to reduce the effect of the PCR amplification during the RNA-Seq library preparation. It results that the duplicated reads of a same fragment will only count once.

Concerning the downstream analysis, a supplemental challenging step is to remove barcodes corresponding to "doublet cells". Doublet cells, or doublets, are barcodes for which two cells have been taken together. Practically, it occurs when two cells were put in the same droplet or the same well on a plate, depending on the single-cell technology. The detection of doublets is not an easy task *in silico*. The library size of doublet is expected to be higher, but it does not represent a reliable criterion to remove barcodes. Recently, several advanced methods have been developed to improve the detection of doublets. One of these methods allows adding different tags to all the cells and they are analysed separately to define droplets/well containing singlets or doublets [116]. Another approach relies on the combination of several individuals inside a multiplexed SC experiment and it can associate each cell to the correct genotype using natural genetic variation [117]. Moreover, efforts are also made to computationally detect them [118].

It relies on the identification of a transcriptionally aberrant signature (e.g. mixing two distinct cell type signatures).

The methods for most of the downstream analyses of single-cell data are similar to those for bulk data. However, the normalization step is particularly affected by the "single-cell" approach. Firstly, the number of genes detected by cell is quite low (it varies with the technology used) and the variable efficiency of the RNA capture (~10-30%) between all cells leads to a "sparse" gene count matrix with many zero-count values. None of the current pipelines present an ideal method for this step as it relies on the evolution of the SC technology in the future. For example, the Seurat R package from Satija Lab started by employing a global-scaling normalization method called "LogNormalize" that normalizes the gene expression values for each cell by the total expression, multiplies this by a scale factor, and log-transforms the result [119, 120]. Their effort in improving this step leads to a new model called "SCTransform" which relies on the use of a regularized negative binomial regression for normalization and variance stabilization [121].

Finally, the visualization of a SC dataset also implies some computational adaptation. An initial linear dimensional reduction of the data is performed through PCA and the identification of different clusters of cells is obtained through different clustering algorithms using the PCA space. In recent years, several non-linear dimensional reduction techniques, such as **UMAP** (Uniform Manifold Approximation and Projection) and **t-SNE** (t-distributed Stochastic Neighbor Embedding), have been adapted for the visualisation of single-cell data [122, 123]. These algorithms learn the underlying manifold of the data to place similar cells together in a low-dimensional space. For an optimal visualization, one should use similar PCs as input for both the clustering and UMAP/tSNE methods (Figure 19).

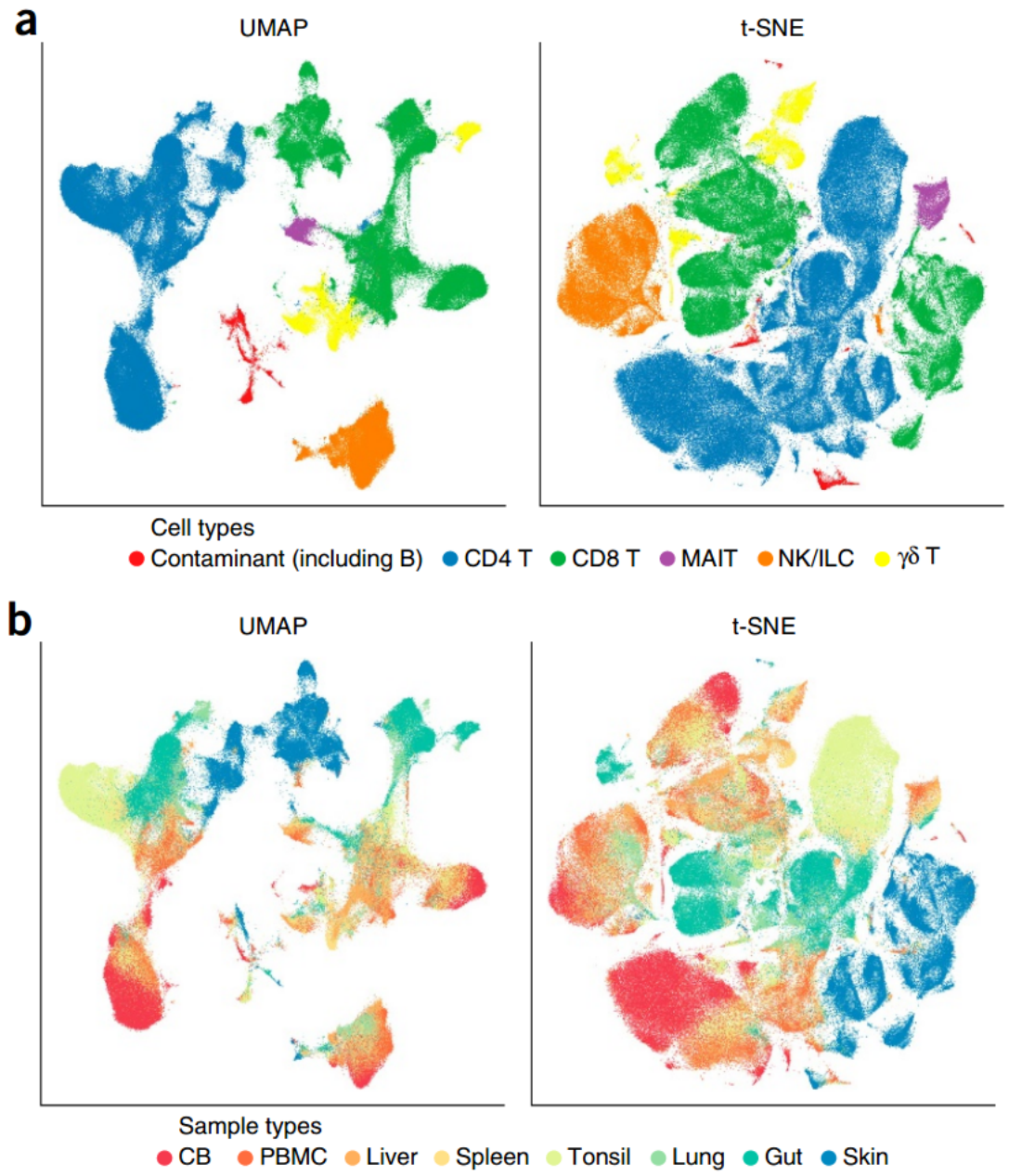


Figure 19: **UMAP and t-SNE (from Becht et al, 2019)**. These non-linear dimensional reduction methods are used to distribute cells into a two-dimensional embedding. Cells are placed according to their similarity. (a) Colored according to broad cell lineages. (b) Colored according to the tissue of origin.

State of the art

1 Transcriptome of adult pancreatic cells

Transcriptomic profiling of cell types is an important tool in modern molecular biology. During nearly 20 years from the early 90s to 2010, microarray analysis and quantitative RT-PCR were used to estimate expression level of chosen gene subsets. Then, NGS technology strongly contributed to the quick development of transcriptomic analysis in the last decade by allowing one to work more precisely and accurately at the genome-wide level. RNA-Seq protocols have constantly evolved to be more sensitive for different types of research. However, the bulk RNA-Seq approach had his limit of application when the ability to isolate cell populations prior to RNA extraction was inefficient. When studying animal models, the use of transgenic lines expressing fluorescent proteins in specific cell types has allowed good cell population isolation using FACS. Using human tissue, this approach was not possible and specific cells of interest were difficult to retrieve efficiently. Fortunately, the recent emergence of scRNA-Seq allowed overcoming this barrier and many transcriptomic studies performed on different human tissues have been undertaken recently using this technique.

Pancreas-associated transcriptomic studies have followed the same chronology. Human pancreatic cell transcriptomes have first been studied by microarrays and then bulk RNA-seq using whole pancreatic islets [124] or were focused on beta cells [125, 126]. “Beta cell regeneration” has been one of the driving factors for pancreatic transcriptomic interest. Identification of beta cell specific transcription factors was an important step for programming pluripotent stem cells into beta cells. Beta cell interest implied that the other pancreatic cell types have been somewhat neglected until the emergence of microarray analyses for most of pancreatic cells [128], soon followed by RNA-Seq experiments [129–133]. For human tissues, the use of cell-surface markers for the isolation of specific cell types still presented limitations as it was a source of variability in the sample purity. It is finally with scRNA-Seq that multiple studies using human tissues provided the transcriptomes for most of the major pancreatic cell types [134–138]. However, different technologies were developed leading to different impacts on outcomes.

Limitations of human experimentation led pancreas studies to be transposed to animal models. Mouse represents the most common mammalian model and pancreatic researches make no exception. Pancreatic development or pancreatic diseases such as diabetes or pancreatic cancer have been widely studied in the murine model. Concerning transcriptomic studies, both bulk RNA-Seq and scRNA-Seq experiments have been performed on pancreatic tissues [131, 134, 140, 141]. Zebrafish is another model commonly used to study pancreatic development or diseases. As previously mentioned, genome engineering features of zebrafish led to establishment of multiple transgenic lines in recent years. The specific expression of fluorescent proteins in pancreatic cells in different transgenic models, associated with the transparency of the zebrafish embryo and to its external development, has been a powerful approach to characterize the dynamics of pancreatic development in zebrafish [142–145] . Moreover, relatively facilitated mutagenesis led to the generation of many functional analyses of genes and to the modeling of pancreatic diseases. Several transcriptomic studies have been published in recent years using RNA-Seq or scRNA-Seq [146–148].

The following sections summarize the main function of each endocrine pancreatic cell type as well as some key factors involved in their identity and physiological roles. As the present study is focused on the transcriptome of pancreatic cells, some key transcript markers will be presented.

1.1 Beta cells

Due to their role in both type I and type II diabetes, beta cells have been the most studied pancreatic islet cell types. Many microarray or sequencing-based transcriptomic studies have been performed on beta cells in both healthy or pathologic conditions making the identity of the mature beta cell the most elucidated among pancreatic cells. The maturation of beta cells has been associated with the expression of *UCN3* [149] while many transcription factors involved during endocrine cell differentiation remain highly expressed in mature beta cells. Cooperation between INSM1, NEUROD1 and FOXA2 has been described to act on regulatory sequences to consolidate the maturation program and the disruption of this association leads to an immature phenotype in beta cells [150]. PDX1, NGN3 and MAFA have been described as critical TFs for beta cell development

[151] with PDX1 also associated with beta fate promotion and alpha fate repression [152]. Moreover, PDX1 plays a key role in beta cell survival by maintaining an adequate pool of healthy Î²-cells in adults but the underlying mechanism is still not fully understood [153, 154]. Transcription factors RFX6, NKX6.1 and PAX6 have also been associated with maintaining the beta cell identity, by repressing genes of other cell types or by maintaining functionality [155–157]. Recently, LDB1 (LIM domain-binding protein 1) and its binding partner ISL1 (Insulin Gene Enhancer Protein ISL-1) have also been shown to maintain a mature beta cell state [68, 158].

Physiologically, beta cell identity is mainly characterized by the ability to sense high glycemia and secreting insulin in response, called **G**lucose **S**timulated **I**nsulin **S**ecretion (GSIS). Activation of this mechanism is a criterion that the cells have reached maturity. Biochemical properties underlying this key function have been reviewed by Dr Guy Rutter and his collaborators [56]. Firstly, it has been established that the glucose uptake itself by the cell together with its subsequent metabolism trigger insulin secretion [159, 160]. Secondly, this process is also controlled by many other biochemical regulators which can independently trigger insulin response (secretagogues), or they can modulate sensitivity of insulin release (potentiators) [161]. The « GSIS » response relies on the expression of key genes highly expressed in beta cells, notably at the transcriptomic level. Beside the insulin coding gene *INS*, naturally strongly represented in beta transcriptome, *SLC2A2* (coding for GLUT2 glucose transporter) and *GLP1R* (coding receptor of incretin hormone GLP1) are also highly expressed in beta cells. The same observation is done for *GCK* gene coding glucokinase as glucose phosphorylation is a key step during GSIS and it has been described to play a key role in insulin secretion [64, 162, 163]. Many critical intermediates and cofactors, called metabolic coupling factors (MCFs), are involved in the metabolic signaling of the insulin secretion [164]. It implies that many signaling pathways regulate the beta cells activity. For example, the transcription factor TCF7L2 has recently been associated with PI3K/AKT signaling to regulate the function of beta cell [165]. Calcineurin/NFAT signaling, through CNB1 (*PPP3R1*), has also been associated with regulation of growth and function of beta cells [166]. In recent years beta cell physiology, often associated to their identity, has been vastly investigated. For comprehensive reviews, see [56, 164, 167].

Finally, recent publications have highlighted the functional heterogeneity in pancreatic beta cells population by identifying a hierarchical network of activation during GSIS [168]. Specific beta cells, called “hub” (or “leader”) cells, initiate the response to glucose before triggering and synchronizing response of “follower” cells [169]. Transcriptomic signature underlying “hub” phenotype is still unclear but different expression levels of important genes such as *GCK*, *PDX1* or *NKX6.1* have been observed. Hub cells are representing a metabolically adapted subpopulation of beta cells which also show features of immature cells [168]. Interestingly, a heterogeneity within beta cells has also been described recently for the islet growth in zebrafish [170, 171]

1.2 Alpha cells

Interest in glucagon-producing alpha cells, like the other endocrine cells, has been relatively neglected compared to that of beta cells. For example, a lot is known about the consequences following complete loss of beta cell mass but what happens if we lose another endocrine cell type is relatively unknown [44]. The group of Herrera PL performed the ablation of alpha cells in adult mice and they observed normal activity of both the glucagon signaling and beta cell function [172]. Nevertheless, these non-beta cells have recently been investigated for their potential ability to transdifferentiate into beta cells and some studies have already described this mechanism using alpha cells [173–177]. Moreover, glucagon is secreted in response to hypoglycemia and an altered response of alpha cells in pathophysiology of T2D has already been identified and is still currently investigated [178–181]. Gromada and colleagues reviewed that in addition to hypoglycemia, glucagon secretion is regulated by several paracrine factors [182].

From the transcriptomic perspective, recent outcomes from RNA-Seq and scRNA-seq experiments have allowed to obtain insight of the transcriptome of alpha endocrine cells [128–138, 140, 141]. Beside *GCG* expression, different transcription factors have been identified with a conserved expression in mature alpha cells, such as *ARX*, *IRX1* and *IRX2* genes. The role of NEC2/PCSK2 in the conversion of pro-glucagon to glucagon explains high expression of the gene in alpha cells [183]. *DPP4* gene expression is also highly enriched in alpha cells and encodes the dipeptidyl peptidase-4 responsible on glucagon

degradation [184].

1.3 Delta cells

Less is known about this endocrine cell type. Beside *SST* gene expression, *HHEX* expression has been described to maintain delta cell fate [185]. Chera and colleagues also highlighted delta cell potential to transdifferentiate into beta cell [174]. Enriched expression of many receptors has also been described with notably the expression of leptin receptor (*LEPR*) and ghrelin receptor (*GHSR*) [136].

1.4 Gamma cells

Also called PP cells or F-cells, gamma cells are characterized by the expression of pancreatic polypeptide (PP) and represent a very small portion of islet cell population in human [186]. PP is secreted in response to food intake and has been described as a satiety hormone which also acts as an inhibitor or the glucagon release [187]. Recent scRNA-seq experiments highlighted the expression of pan-endocrine transcription factors but also the expression of genes usually associated to neuronal cells [135], which is coherent with previous observations about the role of nervous control on pancreatic polypeptide release [188, 189].

1.5 Epsilon cells

These ghrelin expressing cells have recently been more investigated [190]. They are mainly found in the gut but they are also present in a very low amount inside islet cell population. Ghrelin expression rises during fasting and has been described as an inhibitor of the insulin secretion in both human and rodents [191–193]. Transcriptome of epsilon cells is nearly completely unknown; the recent scRNAseq experiments gave only few insights of it due to the scarity of these cells. Interestingly, epsilon cells do not express *PAX6* gene and recent studies in mice highlighted that *PAX6* inactivation in mature beta cells led to loss of *INS* expression associated to a strong increase of *GHRL* expression [156, 194].

2 Pancreatic Development in zebrafish

The next sections describe current knowledge of the pancreatic differentiation program in zebrafish. References are presented for mecanisms that have also been described to work similarly in humans or mice.

2.1 Organ morphogenesis

The pancreas emerges from the endoderm through the development of two different buds, called dorsal and ventral buds, and this characteristic is conserved among vertebrates. In zebrafish, the two buds do not appear at the same time with the dorsal bud already starting to rise at 24 hpf (hours post fertilization). Moreover, the dorsal bud in zebrafish will only contribute to form the initial pool of endocrine cells. This pool will give rise ti a large part of the **primary islet**, a major principal islet of Langerhans found in zebrafish [195]. The ventral bud, emerging at 40 hpf, will develop and differentiate into both exocrine tissues (acini and ducts) and into the secondary pool of endocrine cells which will be scattered in pancreas [196]. The group of Dr Nikolay Ninov (Dresden, Germany) recently described that some cells also contribute to the primary islet [170]. After its emergence, the ventral bud migrates posteriorly so that both buds merge around 52 hpf [197, 198]. The exocrine tissue starts to grow posteriorly and will acquire a glandular morphology with the formation of acini connected to the ductal tree (Figure 20).

The endoderm is specified at the beginning of gastrulation (shield stage, 6 hpf) [199–202] through the action of the transcription factor *sox32*. The endoderm layer expressing *sox17* will relocate along the anterio-posterior axis during gastrulation to form the endodermal rod which will give rise to the embryonic gut [203–207]. The patterning of the endodermal layer along the antero-posterior axis is controlled by a series of signaling pathways including Nodal [201, 202, 204, 205], Wnt [208], Fgf [209], BMP [210–212] retinoic acid [213–215] and Hedhehog [216] during gastrulation (6 to 10 hpf) as well as induction factors released by the mesoderm during somitogenesis stages (10 hpf to 30 hpf) [217–222].

The pancreatic specification starts well ahead of the morphogical constitution of the

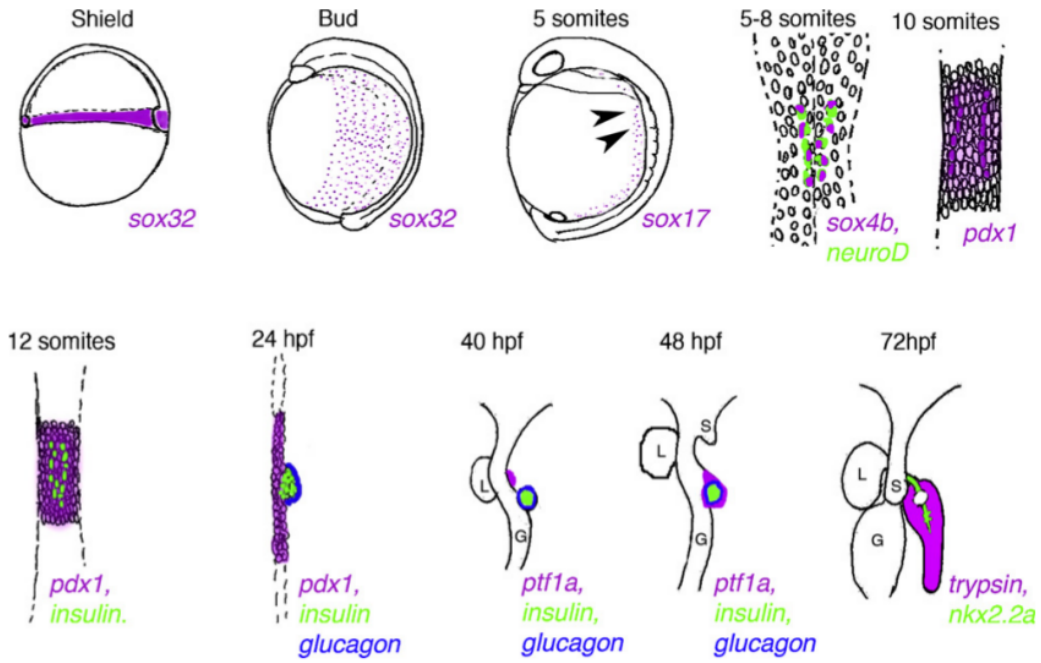


Figure 20: **Schematic representation of zebrafish pancreas development** (from Tiso et al., 2009). Sox4b and neuroD are the first genes expressed in the pancreatic primordium while pdx1 expression defines the future pancreatic region. The dorsal bud emerges at 27 hpf to give rise to first endocrine cells. The ventral bud emerges at 40 hpf and is at the origin of the majority of exocrine pancreas. The two buds merge around 52 hpf.

gland through a complex sequential activation of parallel or interconnected pathways (as reviewed by Tiso and collaborators [223]). This will lead to the creation of a dedicated and precise Pdx1-positive endoderm region called the pancreatic primordium.

2.2 Key players in the pancreatic differentiation program

Starting with genetic pathways, FGF signaling plays an important role during pancreatic differentiation, like in liver, as it controls fate decision between ducts and organ tissues through *fgf10* expression [224]. Evidence showed that Notch signaling is also greatly involved during the process and notably through the different outcomes generated by the different Notch ligands [225].

The differentiation process relies on the sequential expression/repression of key genes, quite conserved across vertebrates, which control the cell differentiation and specific endocrine/exocrine fate choices [226]. Some of them are also involved in defining or maintaining only one specific pancreatic cell identity. As previously mentioned, *Pdx1* is essential during embryogenesis for the formation of the complete pancreas in vertebrates but it also crucial later for beta cell differentiation as for maintaining mature beta cell identity and functions in mice [227, 228]. The same can be say for *ptf1a* whose expression contributes to early pancreas specification, but later, its expression becomes restricted to the acinar cells in vertebrates [142, 229, 230].

Three different intermediate stages during cell differentiation have been described: **multipotent progenitors**, **bipotential progenitors** and **endocrine precursors**. In zebrafish, multipotent progenitors only originate from the ventral bud as they can give rise to all pancreatic cell types. These cells are characterized by the expression of ***ptf1a, pdx1, nkx6.1*** and ***sox9b*** regulatory genes. The bipotential progenitors, also restricted to the ventral bud origin, are characterized by the loss of *ptf1a* expression and thus are not able to differentiate in acinar cells but will give rise to endocrine cells or ducts. In zebrafish, endocrine precursors specification is controlled by the expression of ***ascl1b***, a bHLH transcription factor. It is a difference with mice/human in which it has been vastly described to be the role of ***Ngn3*** (Neurogenin-3) expression [231–234]. In both cases, *ascl1b* or *Ngn3* trigger expression of ***neurod1/Neurod1***, also a bHLH transcription factor,

to drive the differentiation of endocrine cells. In zebrafish, the bHLH factors appear first early in the dorsal bud to give rise to the first endocrine cells and then in the bipotential progenitors from the ventral bud for the emergence of secondary endocrine cells [235].

2.2.1 Exocrine fate during differentiation

As already mentioned, multipotent progenitors are able to give rise to all pancreatic cell types and particularly to acinar precursors. The ***ptf1a/nkx6.1*** balance, in which each gene represses the other, performs a major role in determining acinar fate. ***Ptf1a*** represses Nkx6 genes (*nkx6.1/nkx6.2*) to initiate acinar precursors. Nkx6 genes, by repressing *ptf1a*, promote the bipotential progenitor fate for endocrine and duct development [229, 236]. FGF and TGFβ signaling activation and inhibition, respectively, have also been described to be required for the acinar specification [237, 238] while Notch and Wnt signaling have been associated with an acinar cell fate repression [239, 240].

Contrary to what is observed for acinar or endocrine cells, there is not yet a well described ductal precursor stage. Expression of a ductal fate marker, similar to *ptf1a* for acinar precursors or *ascl1b/Ngn3* for endocrine precursors, has not yet been highlighted making it difficult to identify the specification of ductal cells from bipotent precursors [241]. Nevertheless, evidence has highlighted the role of transcription factors, such as *hnf1b*, *hnf6/onecut1* or *meis1/2*, in controlling specific aspects of duct development [242, 243]. Also, the conserved expression of *nkx6.1* and *sox9b* in mature zebrafish ductal cells has recently been highlighted and has been associated to the potential progenitor role of ductal cells during endocrine tissue regeneration and secondary islets development [235, 244]. We recently highlighted novel genes with an enriched expression in ductal transcriptome of adult zebrafish compared to other pancreatic cells (see Results) [146]. Finally, the role of Notch signaling has also recently been identified to be important for ductal fate in liver and pancreas in zebrafish, as the inactivation of two Notch ligand genes, *jag1b* and *jag2b*, leads to complete loss of duct cells in both organs [245].

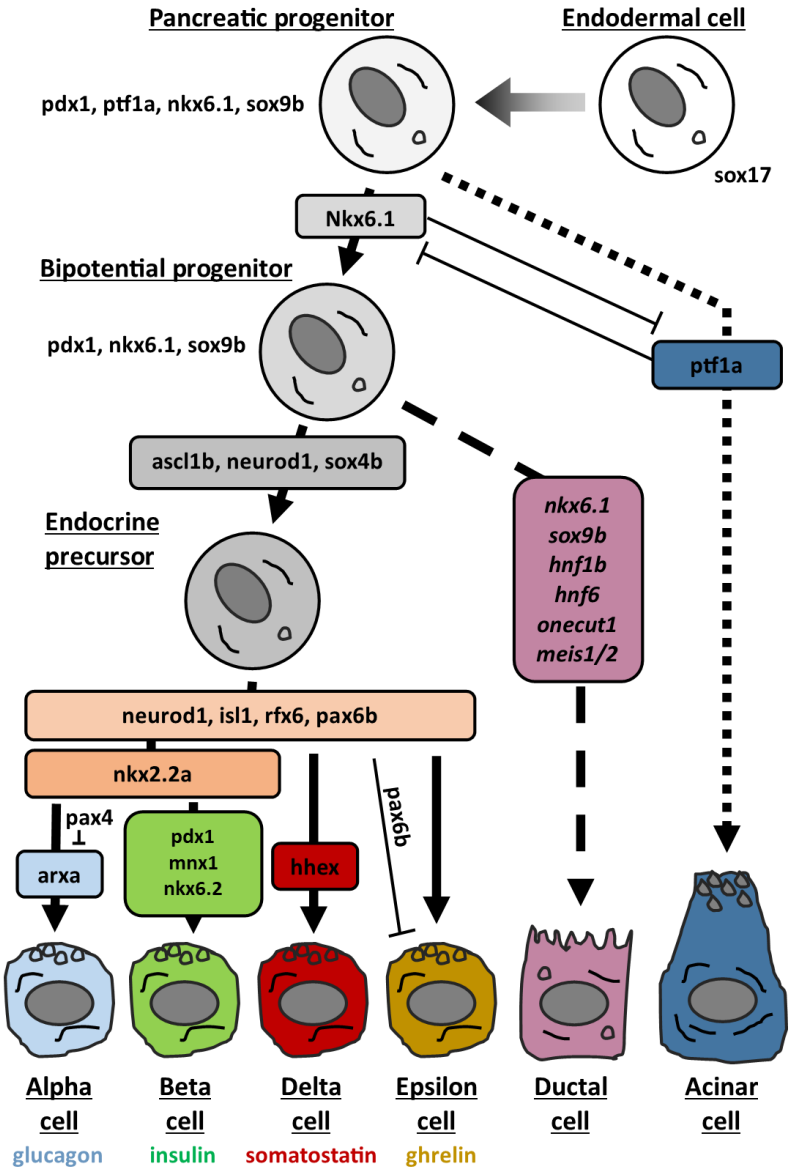


Figure 21: **Differentiation of pancreatic cells in zebrafish**. Pancreatic progenitors are pdx1-expressing cells derived from endoderm region. They are also characterized by the expression of *ptf1a, nkx6.1* and *sox9b*. Mutual repression of *ptf1a* and *nkx6.1* defines the separation between acinar cell and bipotential progenitor fate. *Ascl1b* is the first initiator of endocrine precursor specification. Conversely, the expression of *nkx6.1* and *sox9b* is maintained in ductal cell fate. Endocrine cell subtype differentiation is characterized by the expression of pan-endocrine genes and the expression of genes specific to each endocrine cell subtype.

2.2.2 From endocrine precursors to all different endocrine populations

It has been mentioned previously that in zebrafish, ***ascl1b*** is the first initiator of the endocrine precursors from progenitors (in the dorsal and ventral buds). Quickly, it triggers ***neurod1*** expression and together they induce the specification of endocrine fate [231]. The expression of ***sox4b*** starts nearly concomitantly in the *ascl1b* expression domain and is restricted to endocrine precursors. Interestingly, *sox4b* expression is strongly dependent on *ascl1b* while evidence have also shown the potential of *neurod1* to active *sox4b* expression. Indeed, in *ascl1b* morphants the expression of *sox4b* can be found inside the *neurod1* expression domain [231]. Expression of *ascl1b* and *sox4b* is transient during endocrine differentiation as it is restricted to endocrine precursors while *neurod1* expression is maintained in first hormone expressing cells and in differentiated endocrine cells [231, 246]. Finally, it has also been reported that ***nkx6.1/nkx6.2*** play key role in early stage of endocrine development in zebrafish [247]. As they are already present in early pancreatic progenitors, they have been described to act on both the establishment and the size of the endocrine precursor pool. Only *nkx6.2* expression is then maintained in endocrine cells and progressively restricted to beta cells. Finally, the synergy between nkx6 genes has been suggested to control both alpha and beta differentiation [247].

After the initiation of endocrine fate by these factors, a complex network of transcription factors is subsequentially activated to control and establish the formation of the different endocrine cell type populations. Many of them described during endocrine development in mice/human have been later associated with quite similar roles in zebrafish, such as ***pdx1, nkx2.2a, isl1, pax4, rfx6, mnx1, arxa*** or ***pax6b*** [195, 248–251].

Evidence has shown that ***pdx1*** and ***mnx1*** both cooperate to establish the beta cell fate in zebrafish [251, 252]. As previously mentioned, *pdx1* is required and already expressed in early pancreatic progenitors that give rise to all pancreatic cells. Its “endocrine” expression is progressively restricted to beta cells to establish this fate and to later maintain beta identity in adults [151–153]. Nevertheless, lower *pdx1* expression has also recently been detected in adult ductal cells [146]. The role of *mnx1* in zebrafish slightly differs from that in mice as it is not required in early stages but it is only essential for the beta cell fate [197]. Interestingly, Mnx1 deficiency during development modifies the fate of “pre-beta”

cells towards an alpha cell fate [251]. In zebrafish, the alpha cell fate is determined by the action of ***Arxa*** which is antagonized by ***Pax4***, another homeodomain transcription factor. In mice, *Pax4* has been reported to initiate the beta cell differentiation through the repression of *Arx* gene [253–256]. A slighty different mechanism is observed in zebrafish as *pax4* also represses *arxa* expression to modulate the alpha cell generation but does not participate to beta cell differentiation [248].

Other transcription factors, such as ***nkx2.2a, rfx6***,***isl1*** and **hhex**, are involved in the differentiation of several endocrine cell types and their actions have been maintained from fish to mammals while some slight differences can be noticed between species. Expression of ***Nkx2.2*** in mice starts early in endocrine precursors of the dorsal bud and is later restricted to alpha, beta and gamma cells [257]. It acts on late stages of differentiation/maturation as the inactivation of *Nkx2.2* expression leads to impaired endocrine cell functions without affecting their respective specification [258, 259]. Similar observations have been described for *nkx2.2a* in zebrafish but additional expression in ductal cells has been reported [260]. ***Isl1*** has no influence on the specification of endocrine precursors. However its inactivation affects the generation of hormone producing cells [261]. Recent evidence also highlighted in mice the role of ISL1 and its LIM partner **LDB1** in maintaining the terminally differentially state of endocrine cells [158]. ***Rfx6*** expression in zebrafish starts in early *pdx1*-positive cells of the pancreatic buds and is soon restricted to the endocrine lineage [249]. Finally, the role of **hhex** in the specification of delta cells has recently been identified by the group of Stainier D [262]. The role of ***Pax6*** during endocrine pancreatic differentiation will be discussed in the next section.

3 Transcription factor Paired box 6 (Pax6)

3.1 Discovery and molecular description

The discovery of the DNA-binding transcription factor PAX6 was made in the early 90s as part of an effort to highlight potential genes causing eye diseases in humans and mice. Since then, Pax6 expression has been characterized in **eye, central nervous system** and in **endocrine pancreas**, and 3 decades of PAX6 studies in various research fields have provided insights about multiple PAX6 features [263]. PAX6 is labelled as a "master" regulatory protein, just like GATA1, GATA2, SOX9 and others, which acts as a molecular switch that controls specification/differentiation programs in a gene dosage manner, such as haploinsufficiency can already generate broad spectrum of Aniridia forms and other eyes defects.

PAX6 protein sequence and structure are highly conserved between species [248, 264] (Figure 22). In human, PAX6 exons/introns assembly occupies 20 kilobases (kb) of genomic region but studies have shown that PAX6 regulatory regions extend to 450 kb within the chromosome 11, with mice presenting a syntenic region on chromosome 2 [145, 265, 266]. The two 5'-promoters, called P0 and P1, and the internal promoter (Pα) drive the synthesis of different transcripts which are alternatively spliced leading to the production of the predominant canonical Pax6 (p46) and Pax6(5a) (p48) proteins [267]. Two long non-coding RNAs (lncRNAs) sequences are present in the region, Paupar and Pax6AS1 in humans. The first has been described to be expressed in the brain associated notably with a control on the expression of Pax6 [268].

In zebrafish, the partial duplication of the genome has given rise to two pax6 paralogs, *pax6a* and *pax6b* [269], associated with a subfunctionalization of their original ancestral functions through loss and retention of specific cis-regulatory elements [270, 271] (Figure 23). Both paralogs are expressed during eye and brain development, but only *pax6b* is expressed during endocrine pancreas development [145]. This is due to the retention of a pancreatic regulatory element located upstream from the P0 promoter of the zebrafish *pax6b* gene [145, 272].

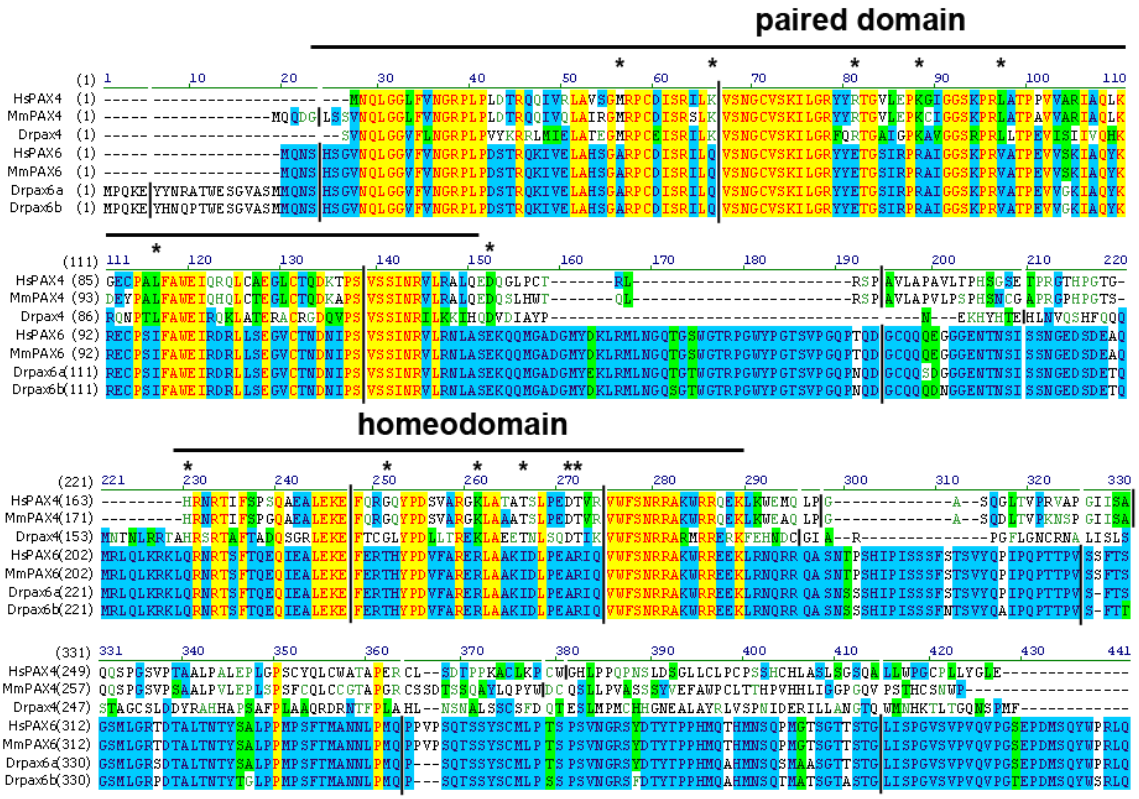


Figure 22: (A) **Alignment of the protein sequence of PAX6 and PAX4 between human, mouse and zebrafish (Djiotsa et al, 2012)**. Regions in yellow represent residues conserved in all sequences while regions in blue represent residues conserved in PAX6 sequences. Locations of PD and HD are displayed. **Hs**: *Homo sapiens*, **Mm**: *Mus musculus*, **Dr**: *Danio rerio.*

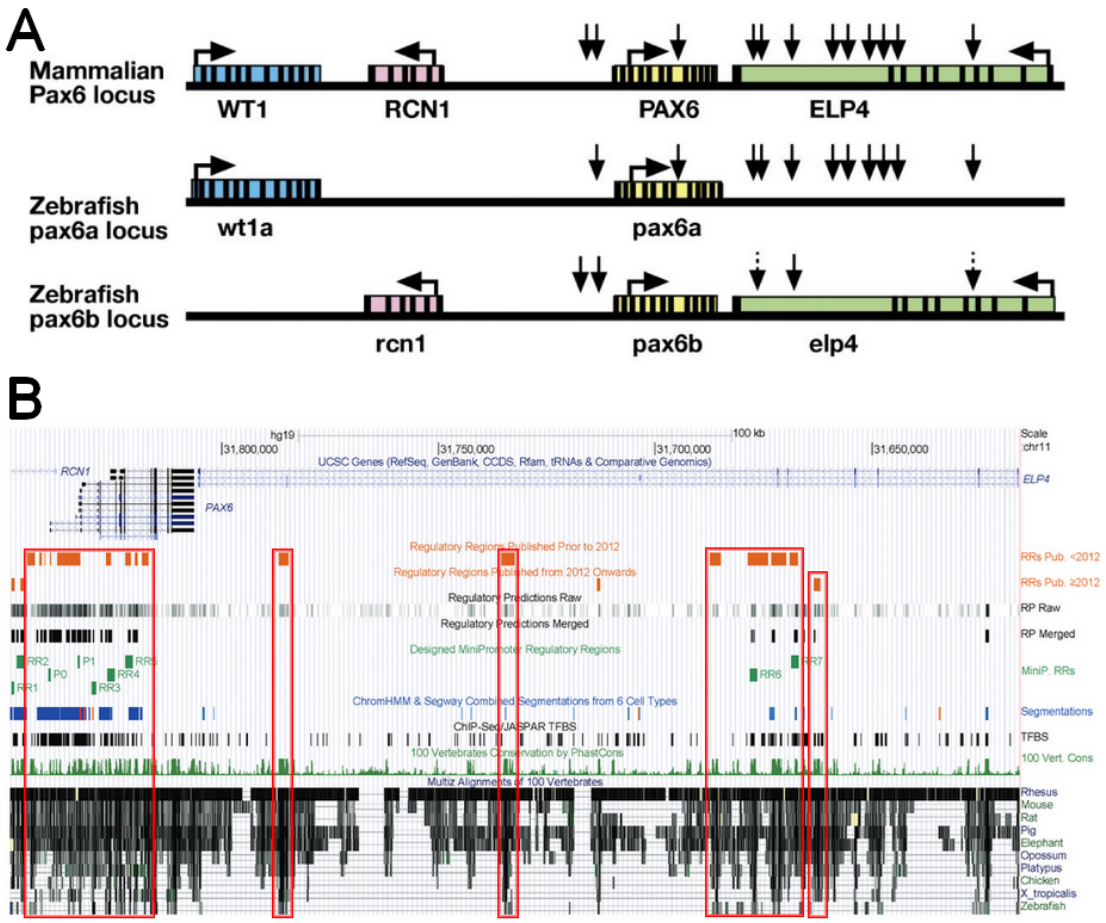


Figure 23: (A) **The PAX6 duplication in the zebrafish genome led to the distribution of its neighboring regions (Kleinjan et al, 2008)** The *pax6a* paralog retained synteny with *wt1a* and with most of the *pax6* cis-regulatory elements (vertical arrows) while the *pax6b* paralog conserved the relationship with *elp4* and *rcn1* but with less regulatory elements. Dotted arrows are for the partially conserved elements. (B) **The PAX6 regulatory regions previously identified and the bioinformatic prediction of putative regulatory regions (Hickmott et al, 2016)**. The alignment of the PAX6 locus across multiple genomes of vertebrates highlights the conserved regulatory regions (red rectangles). These regions present a high sequence similarity between species (10 selected species).

3.2 DNA recognition by PAX6 protein

PAX6 protein contains two DNA-binding domains, a paired domain (PD) and an homeodomain (HD), which are highly conserved between vertebrates. The main Pax6 trancript (p46) encodes both intact domains while the Pax6(5a) alternative spliced mRNA includes exon 5a which adds 14 aa residues and alters DNA-binding properties of the PD [273, 274]. Moreover, PAX6 protein can recognize DNA using both domains together or separately [144, 275], and, combined with alternative splicing, it allows the possibility to bind a large spectrum of DNA-binding sequences. The recent identification of DNA sequences bound by Pax6 *in vivo* during lens and forebrain development using ChIP-Seq has greatly contributed to identify direct target genes of PAX6 [263, 276]. Similarly, Swisa and collaborators identified *Pax6* genomic binding sites in Min6 (mouse insulinoma) cells [156]. These studies resulted to the identification of sets of comparable recognition and consensus sequences (Figures 24). *Pax6* seems to act either as an activator or repressor and the hypothesis has been raised that different DNA-binding sequences can have an influence on this opposite regulation [156]. In zebrafish, evidence has highlighted the dispensable role of *pax6b* homeodomain during pancreatic endocrine differentiation [144].

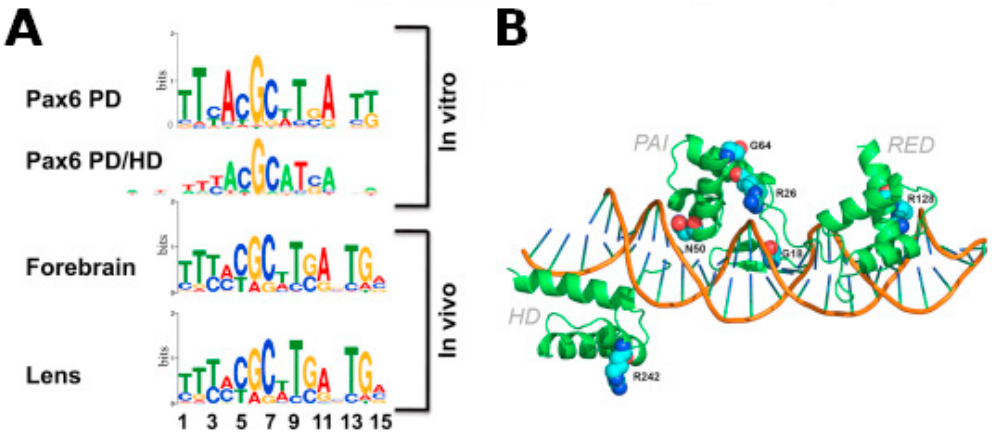


Figure 24: **Pax6 binding to DNA (Cvekl and Callaerts, 2016)**. (A) Pax6 consensus motifs previously identified from *in vivo* and *in vitro* experiments. (B) 3D-structural model of DNA Pax6-binding.

3.3 Pax6b expression during pancreatic endocrine differentiation

The expression of *pax6b* appears in endocrine precursors and it is maintained in both differentiating and mature endocrine cells. In mice and human, PAX6 has been reported to be expressed in all pancreatic endocrine cell types [277–280]. It has been described that *Pax6* exerts direct control on both glucagon and insulin expression through the direct binding of their respective promoter regions [281–283]. *Pax6* is involved in the repression of the epsilon fate [280]. Indeed, in both mice and zebrafish, the inactivation of *Pax6/pax6b* leads to the loss of alpha and beta cells (they are differently affected in both species) associated with an increased number of ghrelin-expressing cells [144, 257]. The loss/decreased expression of insulin, glucagon, somatostatin is concomitant to an increase of ghrelin expression in the islet. The involved mechanisms remain relatively unknown with notably only little information on the genes regulated by Pax6 during endocrine cell differentiation. However, recent studies have described that the inactivation of *Pax6* in adult beta cells leads also to the loss of insulin expression and a strong increased of ghrelin expression [156, 194]. A transcriptomic analysis of these adult beta cells after Pax6 inactivation revealed a large set of PAX6 regulated genes, many of them being involved in beta cell identity and glucose regulation.

In contrast, the PAX6-dependent **Gene Regulatory Network** (**GRN**) has not been investigated during pancreatic endocrine differentiation in mice or zebrafish embryos. PAX6 GRNs have been studied in mice during lens and forebrain development with the identification of direct *Pax6* targeted genes in both tissues [276]. This study revealed that about 70% of PAX6 DNA-binding regions overlap between both tissues while a set of 13 genes are directly regulated by Pax6 in both tissues; this group includes ***Aldh1a3/Raldh3, Ccnd1, Dusp6, Epbh1, Foxp2, Lmo1, Mab21l1, Mrps28, Pax6, Slc39a11, Spsb4, Tenm2*** and ***Tspan7***.

Finally, *Pax6* has been reported to be associated with metabolic diseases. Indeed, *Pax6* heterozygous mutations are responsible for different degrees of glucose intolerance, are associated with a disruption of GSIS and are involved in the development of early-onset diabetes mellitus [284, 285].

Aims & objectives

In previous sections, I have described the role of the pancreas in the digestive process and the homeostasis of specific systems. This role is made possible by the presence of different cell types which present specific functional properties and we are interested in the transcriptomic signature underlying these properties.
In the context of research on **diabetes mellitus**, an important topic focuses on the *in vitro* differentiation of stem cells into pancreatic endocrine beta-like cells [286–288]. In parallel, the plasticity of endocrine cells, and therefore the transdifferentiation potential, has also been the interest of many studies as a possible source for the generation of new beta cells [174–176, 289]. These efforts are dependent on a better understanding of the complete pancreatic endocrine differentiation mechanisms. I have also described in previous sections that maintaining the functional properties of mature cells is closely linked to their identity. Characterization of mature beta cells *in vivo* is required to determine all features that a "neo-generated" beta cell should possess to be considered equally functional. Finally, with the purpose to identify keys elements for these topics, we are interested in the comparison of observations and results obtained in different species.

In this context, the present work focuses on two main topics:

1 The identification of conserved transcriptomic signatures for all major pancreatic cell types

- ***The definition of the transcriptomic signature of the pancreatic cell types in zebrafish***

  > *"What genes are expressed by each pancreatic cell type in zebrafish ?"*

In order to answer that question, we first analyzed RNA-Seq datasets produced in our laboratory for five majors pancreatic cell types [146, 235]. Then, we improved our observations by analyzing publicly available scRNA-Seq data produced on adult zebrafish

pancreas [147] and we determined transcriptomic profiles for seven pancreatic cell types.

- ***The evaluation of the conservation of the transcriptomic signature through different vertebrates***

  > *"What are the genes underlying the identity of each endocrine cell type across vertebrates ?"*

In parallel, we retrieved publicly available RNA-Seq datasets for both mouse and human pancreatic cell types [129–131, 133, 290, 291]. we defined a global endocrine and exocrine signatures in both species. We were thus able to define sets of orthologous genes presenting conserved endocrine or exocrine enriched expression in all three species. Later, we identified available RNA-seq datasets for mouse and human alpha and beta cells, and for each we defined sets of genes with a conserved enriched expression among all three species. At this point, we published our first results.
Then, the emergence of available scRNA-Seq datasets for pancreatic tissue in human allowed us to extend the analyses to all endocrine and exocrine cell types. For all endocrine cell types, we highlighted sets of genes with a conserved enriched expression in both zebrafish and human. These sets of genes are likely to be responsible of the identity of the associated cell type.

2 The role of Pax6b during pancreatic and intestinal endocrine cells development in zebrafish

- ***The identification of transcriptomic changes in endocrine cells following* pax6b *loss-of-function***

  > *"What happens to the pancreatic endocrine transcriptome when pax6b is not expressed ?"*

In our laboratory, RNA-Seq datasets have been produced for pancreatic endocrine cells isolated from 27 hpf embryos using the transgenic lines *Tg(P0-pax6b:GFP)*, in both wild-type and *pax6b*-mutant (**sa0086** [144]) conditions. The similarity of Pax6b-dependent gene regulatory network was determined by comparing the set of genes affected by *pax6b*

inactivation in enteroendocrine cells and pancreatic cells.

> *"Is this network similar in enteroendocrine cells?"*

To answer that question, the same RNA-seq analysis has been performed from enteroendocrine cells using 4 dpf embryos. We compared pax6b-dependent transcriptional networks obtained in each tissue, and we highlighted a significant overlap between the two sets of genes.

- ***The understanding of the alteration of the pancreatic endocrine differentiation process***

> *"Which pancreatic endocrine cells are affected ? How is cell fate influenced ?"*

This question was tackled by combining our results from transcriptomic signatures of adult pancreatic endocrine cells with our list of transcriptomic changes occuring in the *pax6b* mutant. We observed how each endocrine cell type signature is affected in the *pax6b* mutant.

- ***The identification of the genes directly regulated by Pax6b***

> *"What are the genomic DNA-binding sites of Pax6b ?"*

We performed ChIP-Seq experiments using antibodies raised against *Pax6b* on whole embryos at 27 hpf. We retrieved the DNA-binding sites of *Pax6b* and we identified a set of consensus sequences associated. We classified *pax6b* "peaks" according to their genomic localizations.

> *"Which are the genes near the Pax6b binding sites; are their expression regulated by Pax6b ?"*

We annotated the *pax6b* peaks with the closest genes. To identify potential directly regulated genes, we combined our ChIP-Seq results with the list of genes differentially expressed in *pax6b* mutant. This approach highlighted a set of candidate genes for the initiation of the *pax6b loss-of-function* phenotype.

Materials and Methods

1 Zebrafish RNA-Seq for adult pancreatic cell types

We reanalyzed the previously published RNA-Seq data produced in the laboratory [146], available on EMBL-EBI data repository (https://www.ebi.ac.uk/ena/browser/view/PRJEB10140).

2 Zebrafish RNA-Seq of 27hpf and 4dpf embryos

(*Credit to Estefania Tarifeño-Saldivia and Justine Pirson*)

2.1 Zebrafish transgenic and mutant lines, tissue dissection and purification

We used different zebrafish transgenic and mutant lines: Tg(pax6b:GFP) *ulg515* [145] and pax6b *sa0086* [144]. Enteroendocrine cells (EECs) were isolated by dissecting the gut from about 200 Tg(pax6b:GFP) larvae at 4 dpf. Cell dissociation was next performed by incubation in HBSS 1x supplemented with 100 U/ml collagenase IV and 0.3 U/ml Dispase (Life Technologies) for 10 minutes. Cells were washed in HBSS (Mg 2+ and Ca 2+ free) containing 1% BSA and GFP-expressing EECs were selected by two consecutive FACS purifications, the first in the "yield mode" and the second in "the purity mode", using FACS Aria II. Four replicates of EEC containing about 3000 cells were prepared. Pancreatic endocrine cells (PECs) were also obtained from the Tg(pax6b:GFP) ulg515 line by dissecting the dorsal pancreatic bud from about 200 27-hpf transgenic embryos. FACS selection was performed as described for EECs except that cell dissociation was performed in Tryple Select 1X (Gibco) supplemented with 100 U/ml collagenase IV (Life Technologies) for 5 minutes. For the preparations of EECs and PECs from *pax6b* null mutant embryos, the pax6 *sa0086* line was first crossed with the Tg(pax6b:GFP) line; heterozygous pax6b sa0086 fish harboring the transgene (pax6b:GFP) were inbred to generate homozygous pax6b *sa0086* transgenic embryos which were selected based on the absence or reduction of lens. The isolation of EECs and PECs from pax6b sa0086 homozygotes were performed in triplicates following the same procedure that for the

wild-type larvae. The accuracy of pax6b *sa0086* homozygotes selection was verified after the RNA-seq by checking the presence of the null sa0086 allele in 100% of *pax6b* reads in the mutant samples.

2.2 Library preparation

Each EEC or PEC samples obtained after FACS was directly pelleted by centrifugation and resuspended in 3.5 μl of reaction buffer, lysed by freezing in liquid nitrogen and stored at -80°C according to the Smart-seq2 protocol [292]. cDNA was synthesised and amplified by PCR reaction (13 cycles). Quality of cDNA was verified by 2100 High Sensitivity DNA assay (Agilent technologies). 1 ng cDNA was used for preparing each cDNA library using Nextera-XT DNA kit (Illumina) and sequenced on Hi-seq 2000 to obtain around 60 millions of reads (100 base paired-ends).

3 RNA-seq bioinformatic analyses

After trimming of adaptor sequences, reads were aligned to reference genome GRCz11, from Ensembl release 92 (http://www.ensembl.org/index.html), using STAR v2.5.4b [293]. Gene expression matrix was directly obtained from STAR and we performed all pairwise differential expression (DE) analyses using DESeq2 R package [106]. We defined as gene with enriched expression for a cell type all genes with a significative (FDR <= 0.1) enriched expression between the cell type with the highest expression for the gene compared to the second cell type with the highest expression. Genes which failed this test were evaluated for an endocrine or exocrine compartment specific expression by setting a threshold of 10 times enrichment between both conditions. Genes with a low level of expression (CPM lower than 1) were not considered for cell type enrichment.

4 scRNA-seq datasets

4.1 Zebrafish

Zebrafish single-cell data have been generated by the group of Junker J.P. and Ninov N. in the context of cell lineage tracing in zebrafish and identification of cell type [147]. Sequencing data are available on Gene Expression Omnibus under accession number GSE106121. We retrieved their adult pancreatic-associated datasets.

4.2 Human

Human single-cell data used in this work have been generated by three different groups. Count matrices for each experiment were retrieved from their respective data repositories. Segerstolpe data: E-MTAB-5061 (EMBL-EBI ArrayExpress). Muraro data: GSE85241 (NCBI Gene Expression Omnibus). Baron data: GSE84133 (NCBI Gene Expression Omnibus).

5 scRNA-seq bioinformatic analyses

Analyses of scRNA-Seq data were performed using Seurat R package v3.1.0 (http://satijalab.org/seurat/) [119, 120].

5.1 Filtering of cells

We filtered out cell barcodes with low detection level (less than 1000 UMI counts) and cell barcodes with high mitochondrial representation (> 5%). Clusters were evaluated for the expression of key genes and we selected only clusters which could be associated with a high confidence to a pancreatic endocrine cell type, to acinar cells or to ductal cells.

5.2 Differential expression analysis

Differential expression analysis were performed using Seurat *FindAllMarkers* function. This method compares each cluster against the pool of other clusters to define lists of genes with enrichment for the cluster evaluated, which we will refer to as “DE genes”

in this work. To define “marker genes”, we additionally evaluated for each DE gene if there is a significative enrichment (FDR <= 0.001) between the cell type with the highest expression for the gene compared to the second cell type with the highest expression.

5.3 Datasets integration

Integration of all 3 human single-cell datasets was performed using the method proposed by Seurat pipeline (https://satijalab.org/seurat/v3.1/integration.html) [120].

6 ChIP-Seq library preparation

6.1 Collecting embryos

We collected around 5000 zebrafish 27 hpf WT embryos per replicate (n=2). Embryos were extracted from their chorion using pronase and we removed yolk using **Deyolking buffer** (55mM NaCl, 1.8mM KCl, 1.25 mM NaHCO3, 1x protease inhibitor). Embryos were then washed using **Wash buffer** (100mM NaCl, 3.5 mM Kcl, 2.7 mM CaCl2, 10mM Tris-Hcl pH 8.5, 1x protease inhibitor). Crosslinking was performed by placing embryos in 1.85% PFA solution for 10 minutes at room temperature. We used glycine for quenching PFA activity and embryos were washed multiple times with PBS 1x solution. Embryos were then snap frozen and conserved at -80°C.

6.2 Chromatin preparation

We lysed mechanically embryos in **Cell Lysis Buffer** (10mM Tris-HCl (pH7.5), 10mM NaCl, 0.5% Igepal CA-630, 1x protease inhibitor) using tissue grinder. We centrifugated and we resuspended the nuclei pellet in **Nuclei Lysis Buffer** (50mM Tris-HCl (pH7.5), 10mM EDTA, 1% SDS, 1x protease inhibitor). We added 2 volumes of **IP Dilution Buffer** (16.7mM Tris-HCl (pH7.5), 167mM NaCl, 1.2mM EDTA, 0.01% SDS) to the extracted chromatin. Sonication of the chromatin was performed using Diagenode Bioruptor Pico device (10 cycles of 30s On/ 30s OFF). We added Triton 10% before centrifugation. The supernatent was then collected and stored at -20°C.

6.3 Chromatin immunoprecipitation

Antibodies against Pax6b were purified in the laboratory from rabbit injected with the recombinant Pax6b protein (gift of Frederic Biemar). DiaMag protein A-coated magnetic beads (Diagenode) and antibodies were added to chromatin preparation. Immunoprecipitation was performed by rocking samples overnight at 4°C. We washed multiple times the magnetic beads-antibody-pax6b complexes using **Wash Buffer** solution (50mM Tris-Hcl pH8, 150mM NaCl, 2mM EDTA pH8, 1% Igepal CA-630, 0.5 % Sodium Deoxycholate, 0.1 % SDS). Complexes were then placed in **Elution Buffer** (50mM Tris pH8 , 10mM EDTA, 1% SDS) at 65°C in order to eluate protein-DNA complexes from the beads. Reverse cross-linking was performed by placing samples at 65°C overnight. Samples were treated successively with Rnase A and Proteinase K before the Phenol–chloroform extraction of DNA fragments. The retrieved DNA pellet was then resuspended in Illumina Resuspension Buffer.

6.4 Library preparation

ChIP-Seq libraries were prepared using NEBNext® Ultra™ DNA Library Prep Kit for Illumina, with size selection steps performed using SPRIselect For Next Generation Sequencing (Beckman Coulter).

7 ChIP-seq bioinformatic analyses

We used nf-core chipseq pipeline (https://nf-co.re/chipseq) for the QC and the analysis of our ChIP-Seq samples. *De novo* motif anaylsis were performed using HOMER (http://homer.ucsd.edu/homer/) [294]. GO and Pathways enrichment analysis were performed using fishEnrichR (https://amp.pharm.mssm.edu/FishEnrichr/) [114, 115].

8 *In situ* hybridization

Whole mount *in situ* hybridization and fluorescent *in situ* hybridization (WISH and FISH) were performed as described previously [246]. Stained embryos were mounted in Prolong (Invitrogen) with DAPI and imaged using SP5 confocal microsope (Leica).

Results

1 Transcriptomic landscape of major pancreatic cells across vertebrates

The first part of this thesis refers to our work determining the transcriptomic signature of pancreatic cell types in the adult zebrafish. We decided to investigate the evolution of these signatures in vertebrates by comparing our zebrafish results with mammals using murine and human pancreatic transcriptomic data. This work was initiated by Estefania Tarifeño-Saldivia who produced RNA-Seq datasets for five pancreatic cell types using specific transgenic lines: acinar cells (Tg(ptf1a:GFP)), beta cells (Tg(ins:GFP)), delta cells (Tg(sst2:GFP)), alpha cells (Tg(gcga:GFP)/Tg(ins:NTR-mCherry)) and ductal cells (Tg(nkx6.1:eGFP)). At least three biological replicates were produced for each cell type with one additional for acinar cells.

Using those data, Estefania identified sets of genes with enriched expression in endocrine, acinar or ductal cells. Then, genes enriched in endocrine cells were characterized for their specific expression in one endocrine cell subtype or in two of them. Those bioinformatic analyses led to determination of an atlas of genes with an enriched expression in each cell type of the zebrafish pancreas. The next step was to compare our atlas of genes with what is found in mammals. Our strategy was to look for publicly available RNA-Seq data for pancreatic cells and then to analyze them in the same way as we did for the zebrafish data. In this context, we retrieved different data sets obtained from murine and human pancreatic islets but also from whole pancreas samples. Our first interspecies comparison showed conservation of endocrine-enriched and exocrine-enriched genes across all three species. Then, we also collected alpha cells and beta cells RNA-Seq data for mice and human samples, no data being available for delta cells at that moment. We thus performed our second interspecies comparison on alpha-enriched and beta-enriched genes in all three species.

From this work, we identified many novel markers for each pancreatic cell type and some of them were validated by *in situ* hybridization and by functional experiments. These

data contributed mainly to the article "Transcriptome analysis of pancreatic cells across distant species highlights novel important regulator genes" published in BMC Biology in 2017 [146]. **Credit of the work**: Estefania Tarifeño-Saldivia was the initiator of the project, she generated zebrafish RNA-seq data and performed the bioinformatic analysis of these data. I joined her in the project performing the analysis of mammals datasets and the comparison with zebrafish results. Alice Bernard contributed to the project by generating *myt1* and *cdx4* zebrafish mutants. *In situ* hybridization were performed under the collective effort.

This paper is presented in this section and is the foundation of the ongoing effort including an upgraded bioinformatic approach using a more recent version of the zebrafish genome and several available pancreatic transcriptomic data produced with the recent scRNA-seq technology.

Tarifeño-Saldivia et al. BMC Biology (2017) 15:21
DOI 10.1186/s12915-017-0362-x

BMC Biology

RESEARCH ARTICLE Open Access

Transcriptome analysis of pancreatic cells across distant species highlights novel important regulator genes

Estefania Tarifeño-Saldivia, Arnaud Lavergne, Alice Bernard, Keerthana Padamata, David Bergemann, Marianne L. Voz, Isabelle Manfroid and Bernard Peers*

Abstract

Background: Defining the transcriptome and the genetic pathways of pancreatic cells is of great interest for elucidating the molecular attributes of pancreas disorders such as diabetes and cancer. As the function of the different pancreatic cell types has been maintained during vertebrate evolution, the comparison of their transcriptomes across distant vertebrate species is a means to pinpoint genes under strong evolutionary constraints due to their crucial function, which have therefore preserved their selective expression in these pancreatic cell types.

Results: In this study, RNA-sequencing was performed on pancreatic alpha, beta, and delta endocrine cells as well as the acinar and ductal exocrine cells isolated from adult zebrafish transgenic lines. Comparison of these transcriptomes identified many novel markers, including transcription factors and signaling pathway components, specific for each cell type. By performing interspecies comparisons, we identified hundreds of genes with conserved enriched expression in endocrine and exocrine cells among human, mouse, and zebrafish. This list includes many genes known as crucial for pancreatic cell formation or function, but also pinpoints many factors whose pancreatic function is still unknown. A large set of endocrine-enriched genes can already be detected at early developmental stages as revealed by the transcriptomic profiling of embryonic endocrine cells, indicating a potential role in cell differentiation. The actual involvement of conserved endocrine genes in pancreatic cell differentiation was demonstrated in zebrafish for *myt1b*, whose invalidation leads to a reduction of alpha cells, and for *cdx4*, selectively expressed in endocrine delta cells and crucial for their specification. Intriguingly, comparison of the endocrine alpha and beta cell subtypes from human, mouse, and zebrafish reveals a much lower conservation of the transcriptomic signatures for these two endocrine cell subtypes compared to the signatures of pan-endocrine and exocrine cells. These data suggest that the identity of the alpha and beta cells relies on a few key factors, corroborating numerous examples of inter-conversion between these two endocrine cell subtypes.

Conclusion: This study highlights both evolutionary conserved and species-specific features that will help to unveil universal and fundamental regulatory pathways as well as pathways specific to human and laboratory animal models such as mouse and zebrafish.

Keywords: RNA-seq, Comparative transcriptomics, Pancreas, Endocrine cells, Acinar cells, Ductal cells

* Correspondence: bpeers@ulg.ac.be
Laboratory of Zebrafish Development and Disease Models (ZDDM), GIGA, University of Liège, Avenue de l'Hôpital 1, B34, 4000 Sart Tilman, Liege, Belgium

Background

Pancreas is a vital organ playing crucial function in the metabolism of all vertebrates. Acinar cells, the most abundant cell type of the pancreas, produce the digestive enzymes that are conveyed to the gut by the pancreatic ducts. The pancreatic endocrine cells are grouped in the Langerhans islets and secrete diverse hormones controlling metabolism and glucose homeostasis. Five endocrine cell subtypes (alpha, beta, delta, PP, and epsilon cells) have been described in pancreatic islets, each characterized by the expression of a particular hormone (glucagon, insulin, somatostatin, pancreatic polypeptide, and ghrelin, respectively). Many transcriptomic studies have been focused on pancreatic endocrine cells, and notably on beta cells, due to their implication in the development of diabetes. Microarrays and RNA sequencing (RNA-seq) were conducted on endocrine pancreatic cells isolated from human (healthy or diabetic persons) or rodents at adult and embryonic stages [1–9]. More recently, RNA-seq performed on endocrine cell types isolated by FACS or at single-cell level allowed to define the genes enriched in each endocrine cell type in human and mice [7, 10–21]. However, no comprehensive interspecies comparison has been performed so far to define the conserved signatures for each pancreatic cell type, except for two studies comparing human and murine beta cells and reporting some notable differences between these two mammalian species [12, 21]. Comparison of transcriptomes between species is a straightforward approach to identify tissue-specific or cell type-specific genes playing crucial functions in the physiology of the studied tissue or cell. Indeed, if a gene is essential in a differentiated cell, strong constraints will maintain its expression throughout evolution and its tissue-specific expression will be detected in most species. In accordance to this view, several studies have reported conservation of organ-specific expression for a large set of genes among species [22–24]. Nevertheless, there are also striking divergences in gene expression patterns even between close vertebrate species such as human and mice, which may contribute to physiological adaptations [25, 26]. Comparative studies between evolutionary distant species are useful to identify the set of genes displaying highly conserved cell type-specific expression and likely playing a fundamental function in the studied cells. Such analyses have not been performed yet on pancreatic cells due to the lack of transcriptomic data from pancreatic cells isolated from lower vertebrates such as zebrafish. To tackle this lack of knowledge, we first determined the transcriptomic landscape of the three major endocrine cell types (alpha, beta, and delta cells) as well as of the acinar and ductal cells from zebrafish. Analysis of these zebrafish datasets allowed us to define the signature of each cell type. Then, comparison with published human and murine pancreas data led to a definition of the conserved signatures. Furthermore, by determining the transcriptome of endocrine cells from early stage embryos, we identified genes expressed during endocrine cell differentiation and putatively involved in this process; among them, *myt1b* and *cdx4* are shown to be essential for endocrine cell differentiation in zebrafish. Thus, our list of pancreatic conserved genes represents a useful resource for studies related to pancreatic development and disease such as diabetes and pancreatic cancer.

Results

Transcriptomic profiles of the different pancreatic cell types isolated from adult zebrafish

We purified the different pancreatic cell types from adult zebrafish using a series of transgenic reporter lines allowing the selection of these distinct cells by fluorescence-activated cell sorting (FACS). Acinar cells were obtained from the BAC transgenic lines *Tg(ptf1a:GFP)* [27]. The endocrine beta and delta cells were isolated, respectively, from the transgenic lines *Tg(ins:GFP)*^{ulg021Tg} (see Methods section) and *Tg(sst2:GFP)* [28]; the alpha cells were obtained from the *Tg(gcga:GFP)*/*Tg(ins:NTR-mCherry)* line through selection of GFP^{+}/mCherry cells (as many beta cells were found to express Tg(gcga:GFP) transgene at a lower level, Additional file 1: Figure S1). RNA-seq was performed on three independent preparations for each cell type, except for acinar cells, for which four replicates were prepared. About 60 million of paired-end reads were obtained from each Illumina library, 80% of which mapped to the zebrafish genome. We previously reported the transcriptome of pancreatic ductal cells by using the same procedure on the *Tg(nkx6.1:GFP)*^{ulg004Tg} transgenic line [29], and these data were compared in the present study with endocrine and acinar cell transcriptomes.

Principal component analysis (PCA) of all these pancreatic RNA-seq datasets showed a tight clustering of all replicates for each pancreatic cell type (Fig. 1a), underscoring the high reproducibility of the data. As expected, the PCA also revealed a closer clustering of the three endocrine cell subtypes compared to the ductal and acinar cell types; however, when PCA is performed only with the endocrine datasets, clear distinct transcriptome profiles are observed for the alpha, beta, and delta cell subtypes (Fig. 1b). Comparison of the expression levels of various known markers of each pancreatic cell type confirmed the high purity of each cell preparation. Indeed, *glucagon a* (*gcga*), *insulin* (*ins*), and *somatostatin 2* (*sst2*) were selectively detected at very high levels in alpha, beta, and delta cell libraries, respectively, representing in average 24%, 10%, and 32% of the total reads number in the corresponding libraries, while being detected at much lower levels in the other libraries (Table 1). The *trypsin* (*try*) and *chymotrypsin-like*

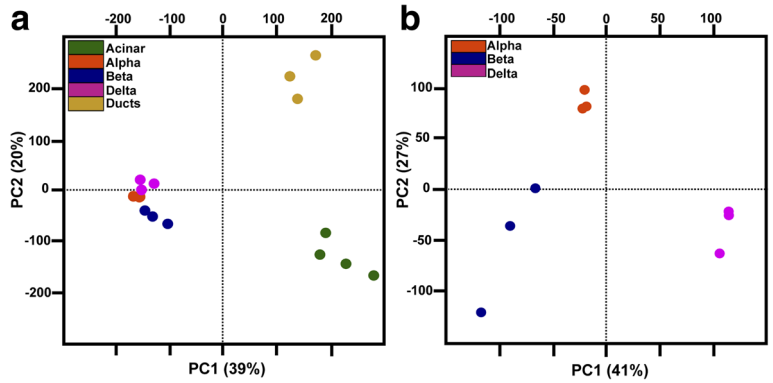


Fig. 1 Global analysis of the zebrafish pancreatic RNA-seq data. **a** Principal component analyses (PCA) of gene VSD (Variance stabilizing transformation) calculated by DESeq package for the 16 zebrafish pancreatic datasets. **b** PCA of gene VSD for beta, alpha, and delta cells (nine samples in total). The PCA plots show a close clustering of all replicates and distinct clusters for each pancreatic cell type. PCAs were calculated using all the 33,726 genes annotated on Zv9 version 75 ensembl

elastase member1 (*cela1*) genes were the strongest expressed genes in acinar cells representing each about 10% of all total reads of the acinar cell libraries. These two acinar markers were not detected at significant levels in endocrine datasets consistent with an accurate cell sorting. Among the genes expressed at highest levels in ducts, we find the *cldnb*, *sdc4*, and the *epcam* genes coding for cell adhesion molecules, each representing less than 1% of total reads of ductal datasets. All these results indicate an accurate and reproducible sorting of the different pancreatic cells allowing the identification of genes selectively expressed in each pancreatic cell type. Expression values for all genes in all samples are shown in Additional file 2: Table S1 and Additional file 3: Tables S2.

Identification of genes enriched in endocrine, acinar, and ductal pancreatic cells

The clear distinct transcriptomic profiles observed for endocrine, acinar, and ductal cells prompt us to identify, in a first step, all genes presenting a differential expression in these three pancreatic tissues (with at least a four-fold enrichment and adjusted $P < 0.05$). By using these cut-off values, 1853, 1430, and 492 genes were found to be enriched in endocrine, ductal, and acinar cells, respectively (Fig. 2a and b, gene lists are given in Additional file 4: Table S3). As expected, the zebrafish endocrine-specific genes include several orthologs of mammalian genes known as endocrine markers such as those involved in hormone regulation and secretion, like the proprotein convertases (*pcsk1* and *pcsk2*), carboxypeptidase E (*cpe*), secretogranins (*scg2a*, *scg3*, and *scg5*), ATP-dependent potassium channels (the *kcnj11* and *Sur1*/*abcc8* subunits), and the voltage-dependent type calcium channels (*cacna1c* and *cacna1da*), among others. Moreover, most of the transcription factors previously shown to be crucial for the differentiation of endocrine cells (*neurod*, *isl1*, *pax6b*, *insm1a*, etc.) [30] are detected in the endocrine signature validating our RNA-seq data. Interestingly, many other transcription factors, whose function in endocrine cells is still not known, are present in this list such as *egr4*, *creb3l1*, *lmo1*, *cdx1b*, or *cdx4* (Fig. 2b). Gene ontology (GO) enrichment analysis using DAVID revealed known biological pathways in endocrine cells such as "potassium ion transport", "regulation of secretion" and "regulation of exocytosis" as well as "response to glucose", and "G-protein coupled receptor signaling" (Fig. 2c). Indeed, many G-couple protein receptors (GPCRs) and several regulators of G-protein signaling, like *gpr12*, *gpr22*, *gpr27*, or *gpr63* as well as *rgs4*, *rgs5a*, *rgs8*, or *rgs17*, among others, were found enriched in the endocrine cells. While some of these regulatory proteins have been previously reported to control pancreatic islet activity [31, 32], others were not yet known to have a selective expression in pancreatic endocrine cells nor to play a role in islet physiology.

Similar observations were done for the acinar and ductal cell transcriptomic signature. Indeed, as expected,

Table 1 Percentage of the reads obtained for highest expressed markers in each type of library

Gene	Alpha	Beta	Delta	Acinar	Duct
gcga	23.8	0.5	0.9	0.0	0.0
ins	0.1	10.1	0.0	0.0	0.0
sst2	0.2	0.1	31.5	0.0	0.0
try	0.1	0.0	0.1	9.8	1.7
CELA1	0.1	0.0	0.1	12.7	1.7
cldnb	0.0	0.0	0.0	0.0	0.5
epcam	0.0	0.0	0.0	0.0	0.8
sdc4	0.0	0.0	0.0	0.0	0.5

Percentages of mapped reads obtained for some genes on the total number of mapped reads in each type of libraries (Numbers are the means of percentages obtained from replicates)

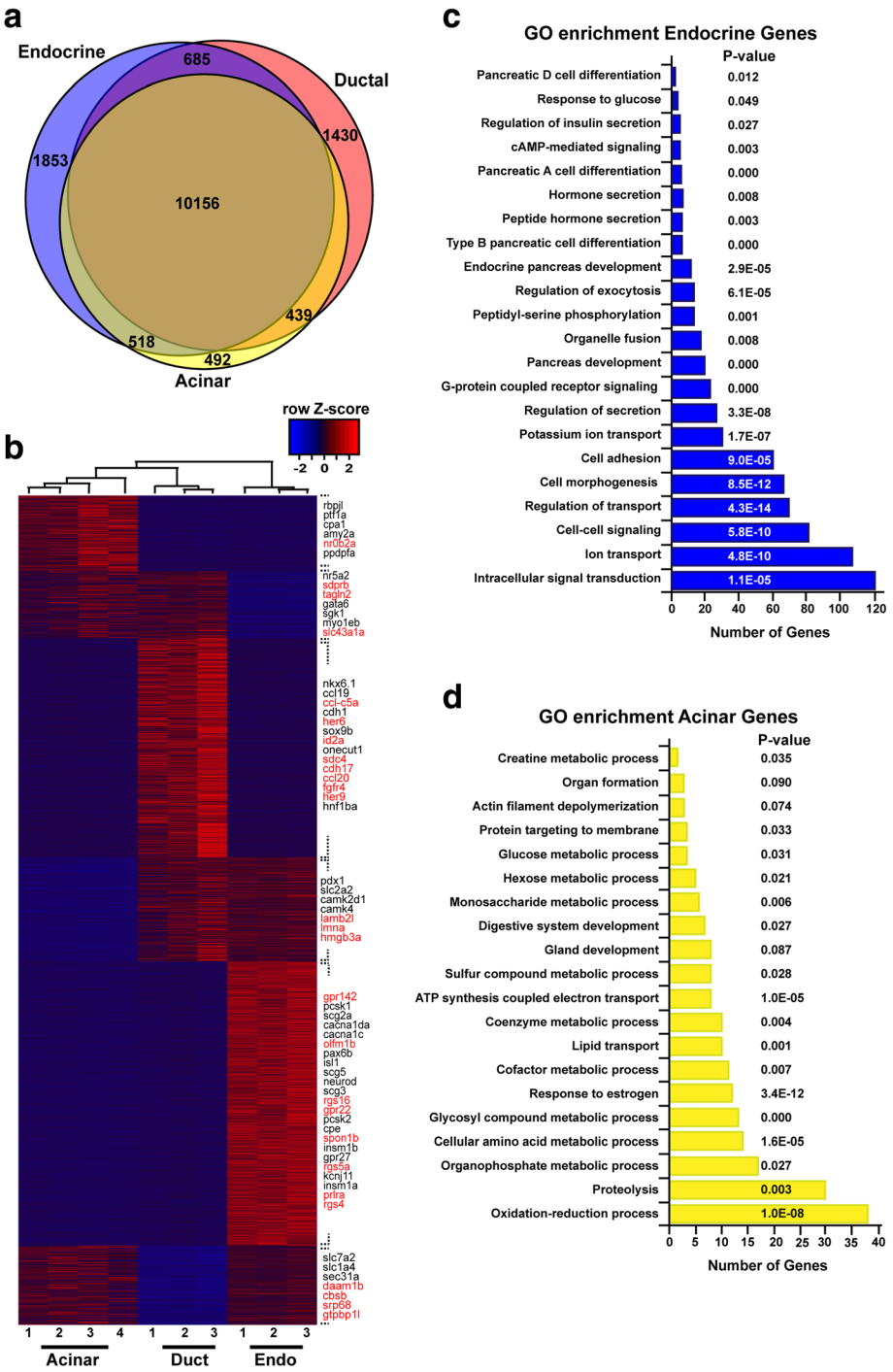


Fig. 2 Transcriptomic signatures of zebrafish endocrine, ducta,l and acinar cells. **a** Venn diagram showing the number of genes with acinar-, endocrine- and ductal-enriched expression obtained with DESeq2 using a cut-off ratio of four-fold and adjusted $P < 0.05$. **b** Heatmap plot showing the expression pattern of all differentially expressed genes. Genes listed on the right side of the plot are examples of either known markers (black) or new genes with endocrine-, acinar-, and ductal-enriched expression discovered by this analysis (red). The three endocrine samples were generated in silico by combining the raw data obtained from alpha, beta, and delta cells as described in the Methods section. **c** and **d** Gene ontology enrichment analysis for the 1853 genes enriched in endocrine cells (**c**) and the 492 genes enriched in acinar cells (**d**) displaying the most enriched biological process

many genes coding for digestive enzymes, such as *trypsin* (*try*), *elastase* (i.e., *ela* and *cela*), *trypsin-like* (*tryl*/*zgc:66382*), *amylase* (*amy2a*) as well as the known acinar transcription factors *ptf1a* and *rbpjl*, were found enriched in acinar cells. The acinar gene list also includes novel markers such as genes encoding for the

transcription factors *esr2a*, *klf15*, or *nr0b2a* (Fig. 2b and Additional file 4: Table S3). As expected, GO enrichment analysis revealed, among the biological processes active in acinar cells, "Gland development" and "Digestive system development", "Cellular amino acid metabolic process", "Proteolysis" as well as "ATP synthesis coupled electron transport" (Fig. 2d). As for the duct transcriptome, the analysis reveals novel markers such as *id2a* and *frzb* in addition to known markers like *sox9b*, *nkx6.1*, *onecut1*, and *ctgfa*, as we recently described [29]. All together, these analyses confirm that many known markers and biological pathways display the same pancreatic enrichment in zebrafish as in mammals, and also highlight many novel cell type-specific genes not previously reported to display such selective expression in mammals. This led us to compare comprehensively the endocrine- and exocrine-enriched genes across zebrafish and mammalian species, thereby defining the conserved specific signatures among vertebrates.

Comparison of the pancreatic endocrine and exocrine transcriptomic signatures across zebrafish, mouse, and human

The evolutionarily conserved expression of a gene in a specific cell type is a strong argument for its crucial function in that cell. This concept is widely supported by numerous studies showing conserved expression of many transcription factors in development and cell differentiation [23, 33], including pancreatic cells [34, 35]. To perform interspecies transcriptome comparison, we retrieved, from public databases, several RNA-seq datasets obtained on murine and human pancreatic islets, whole pancreas [11, 36–38], as well as a human pancreatic sample enriched in acinar tissue [9]. While mouse RNA-seq data are not presently available for acinar enriched preparations, the genes with either endocrine-enriched or exocrine-enriched expression (named hereafter as "endocrine or exocrine genes") can nevertheless be identified by comparing RNA-seq from purified pancreatic islets versus RNA-seq from whole pancreas (composed of more than 90% of exocrine cells). When a global comparison of the zebrafish endocrine and acinar datasets with the murine and human pancreatic data was performed using a PCA plot analysis, we observed that the data clustered according to the tissue type (endocrine and acinar) in the first component (PC1: representing 55% of variance), while the data tended to cluster according to the species in the second component (PC2: representing 22% of variance) (Fig. 3a). Thus, this global analysis suggests the existence of sets of genes displaying an evolutionary conserved endocrine and exocrine expression. Endocrine and exocrine genes were identified from human and murine datasets by selecting all genes presenting at least four-fold higher expression in each pancreatic tissue (with adjusted $P < 0.05$). Next, we compared the endocrine and exocrine gene lists between the three vertebrate species. The Venn diagram on Fig. 3b indicates that, while a large fraction of endocrine-enriched genes are species-specific, some sets of genes display a conserved endocrine-enriched expression in two species or in the three species. Indeed, among the 1853 zebrafish endocrine genes, 251 are also classified as "endocrine" in human and mice. This endocrine signature, conserved between zebrafish, mice and human (named "ZMH"), includes many factors known to be involved in hormone maturation, secretion, and regulation. Many transcription factors crucial for endocrine cell differentiation [30, 39] are found in this ZMH conserved endocrine signature (Table 2), thereby validating this method to identify genes with important physiological function. The ZMH endocrine genes (given in Additional file 5: Table S4) also comprise components of signaling pathways reported to control pancreatic endocrine cells in mammals such as the glutamate receptor *gria2a* and *gria2b* [40], regulators of G-protein signaling (*rgs4* and *rgs16*) [32, 41], *urocortin3* [5, 42], ion channels (*abcc8*, *cacna1*, and *scn1lab*), and calcium dependent proteins (*c2cd4* and *scgn*). This indicates that many regulatory processes controlling endocrine cell development and physiology have been maintained from fish to human. Interestingly, the ZMH endocrine genes include several regulatory genes whose function in pancreatic cells is still unknown or not well defined such as the kinases *map3k15* and *mast1*, the signaling factor *gpr158*, as well as the transcription factors *npas4a*, *lmo1*, *fev*, and *etv1*. Similarly, the comparison of the exocrine-enriched genes across zebrafish, mice, and human shows that a fraction of genes displays conserved exocrine enrichment (Additional file 6: Figure S2) – among the 2361 zebrafish exocrine genes (combined acinar and ductal genes from Fig. 2a), 127 show exocrine enriched expression also in human and mice. This conserved exocrine signature (Additional file 7: Table S5) includes, as expected, many digestive enzymes as well as known exocrine transcription factors such as *ptf1a*, *rbpjl*, and *nr5a2* [36, 43], but also regulatory genes with unknown pancreatic function, like *klf15*.

The interspecies comparison of both exocrine and endocrine signature also indicates that many genes display the same enrichment in only two species (e.g., "HM": between human and mice; "ZM": zebrafish and mouse). As expected, the number of conserved HM endocrine and exocrine genes is higher than the number of conserved ZH and ZM genes. Overall, this global analysis indicates that, while significant divergence is observed between species at the level of endocrine and exocrine signatures, hundreds of genes have nevertheless maintained a common expression pattern, among which are most of the known pancreatic regulatory genes.

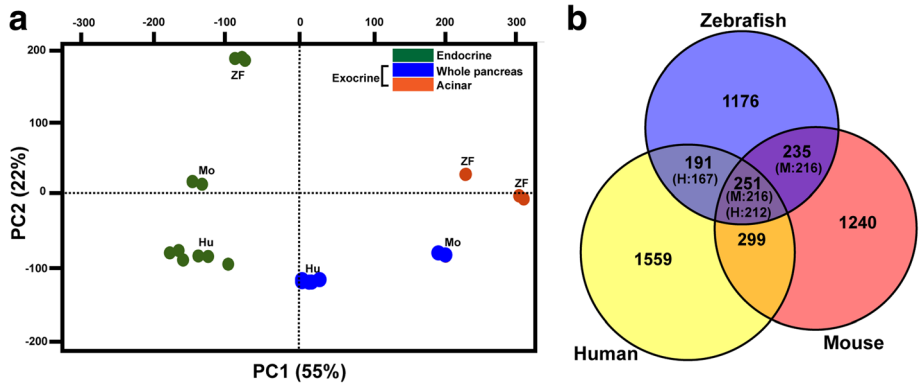


Fig. 3 Conservation of the pancreatic endocrine signature among vertebrates. **a** Principle component analysis (PCA) performed on human and mouse whole pancreas and islet RNA-seq datasets and including the zebrafish endocrine and acinar datasets. The analysis was performed using the 9393 genes displaying 1-1-1 orthology relationship between zebrafish, mouse, and human using a total of 24 RNA-seq samples. The endocrine datasets of zebrafish, mouse, and human cluster along the PC1 axis representing 55% of the variance, indicating a conserved endocrine signature. The human pancreatic sample enriched in acinar tissue [9] clusters with the human whole pancreatic samples [37] due to the very high proportion of acinar cells in pancreas. **b** Venn diagram showing the number of endocrine-enriched genes found only in zebrafish, mouse, or human and those displaying conserved endocrine-enrichment in two species or in the three species (shown in intersections). Due to gene duplications in some species and often in zebrafish, the number of corresponding murine (M) or human (H) orthologous genes is given in brackets in each intersection. The full list of conserved endocrine-enriched genes is given in Additional file 5: Table S4

Identification of genes expressed in embryonic endocrine cells

While many transcription factors controlling pancreatic endocrine cell differentiation in embryos remain expressed and functional in mature cells from adults, there are nevertheless striking differences between the transcriptomes of fetal and adult cells [5]. In order to determine the fraction of the endocrine-enriched genes in adult zebrafish that are already expressed in the first embryonic endocrine cells, we determined the transcriptomic profile of endocrine cells isolated from *Tg(pax6b:GFP)* embryos at 27 hpf. We detected 9919 genes expressed in embryonic pancreatic cells above the threshold level of 100 Normalized count. By comparing these data with adult endocrine data, we found that, among the 1853 genes enriched in adult endocrine cells, 911 (49%) genes were already detected in embryonic cells (full gene list shown in Additional file 8: Table S6). The expression of some of these endocrine-enriched genes was further characterized by in situ hybridization (ISH) on whole zebrafish embryos (Fig. 4). We analyzed genes identified as endocrine enriched in the three species (the "ZMH" genes: *pcsk1*, *pcsk2*, *fev*, *cpe*, *etv1*, *map3k15*, and *lmo1*), in two species (the "ZM" gene *cdx4* and the "ZH" gene *tbx2b*), or only in zebrafish (the "Z" genes *gpr22*, *dkk3b*, *pnoca*, *ppdpfb*, *scinlb*, and *spon1b*). All these genes were detected in the embryonic endocrine pancreatic cells by ISH, validating the RNA-seq data. Thus, these data indicate that the early embryonic endocrine cells from the dorsal pancreatic bud express a large fraction of the genes constituting the adult endocrine signature.

Characterization of the transcriptomic signatures for the endocrine cell subtypes

In order to define the molecular signatures of alpha, beta, and delta cell subtypes, the 1854 endocrine genes identified in adult zebrafish (from Fig. 2a) were classified according to their differential expression in the distinct endocrine cell subtype (above the threshold of four-fold enrichment, adjusted $P < 0.05$); 73, 70 and 192 endocrine genes were found to be enriched in alpha, beta, and delta cells, respectively (Fig. 5a). The heatmap plot in Fig. 5b presents an overview of the expression pattern for all differentially expressed genes (gene list available in Additional file 9: Table S7). PCA using all 1854 endocrine-enriched genes displays a clear discrimination of the alpha, beta, and delta cell transcriptomes (Fig. 5c). GO enrichment analysis performed on beta cell-specific genes identifies "response to organic substrate", "ion transport", or "response to nutrient levels" as enriched pathways among others (Additional file 10: Figure S3A). The gene expressed at the highest level in the beta cells after *insulin* is *ppdpfb*. This gene is the paralog gene of the *ppdpfa* gene (*pancreatic progenitor cell differentiation and proliferation factor a*), which was reported as specifically expressed in acinar cells and controls their differentiation [44]. Interestingly, the paralog *ppdpfb* is specifically expressed in the endocrine pancreas and mainly in beta cells. We verified the expression of *ppdpfb* in 24 hpf zebrafish embryos by fluorescent ISH (FISH) and, according to the RNA-seq data, we confirmed its expression in many beta cells and in a few alpha cells but not in delta cells (Fig. 6a–c). The RNA-seq data also confirmed the selective expression of the

Table 2 List of conserved endocrine-enriched transcription factors

Gene	Endocrine	Acinar	Duct	Conservation
runx1t1	535	2	7	ZHM
isl1	8613	41	76	ZHM
myt1la	96	0	4	ZHM
fev	5884	128	44	ZHM
neurod1	28467	260	300	ZHM
ascl1a	86	0	1	ZHM
rfx6	1517	26	5	ZHM
pax6b	8059	47	23	ZHM
insm1b	1200	6	39	ZHM
npas4a	4497	11	114	ZHM
arxa	1697	38	1	ZHM
myt1a	532	0	17	ZHM
insm1a	3419	10	248	ZHM
lmo1	1425	4	1	ZHM
etv1	284	0	0	ZHM
esrrga	159	8	2	ZM
rfx2	1590	46	253	ZM
sim1a	164	2	1	ZM
mnx1	1368	144	3	ZM
cdx4	144	0	16	ZM
bhlhe41	2984	712	640	ZM
lmx1bb	145	2	18	ZM
znf516	219	3	31	ZM
esr1	1701	10	59	ZH
tbx2b	1246	7	8	ZH
pgr	399	4	15	ZH
nkx3.2	293	0	26	ZH
nr0b1	284	14	8	ZH
sox11b	49	1	10	ZH

Transcription factors enriched in endocrine cells in zebrafish, human and/or mouse. The level of expression in endocrine, acinar, and ductal cells is shown for each gene in the second, third, and fourth column, respectively. The degree of expression pattern conservation is shown in the last column as ZHM: endocrine-enriched in zebrafish, human, and mouse; ZM: endocrine-enriched in zebrafish and mouse; ZH: endocrine-enriched in zebrafish and human

zebrafish *nkx6.2* gene in beta cells, as previously reported [45]. Expression of the zebrafish *pcsk2* gene is more enriched in beta cells compared to the *pcsk1* gene (beta/alpha enrichment of 5.5- and 2.5-fold, respectively), while in mouse and human, only PCSK1 was reported to be more enriched in beta compared to alpha cells [4, 12, 46]. The enrichment of *pcsk2* in beta cells was confirmed by FISH on zebrafish embryos (Fig. 6d–f). The zebrafish estrogen receptor 1, *esr1*, is also strongly enriched in beta cells (10575 normalized counts in beta cells versus 671 in alpha and 71 in delta cells). Finally, the beta cell-selective expression of *spondin 1b* (*spon1b*), an activator of the Wnt pathway, was also validated by FISH on zebrafish embryos (Additional file 11: Figure S4).

As expected, the genes specifically expressed in alpha cells include those coding for the two zebrafish glucagon hormones (*gcga* and *gcgb*) as well as for the Arx transcription factor (*arxa*) but also novel markers including peptide hormones such as *prepronociceptin* (*pnoca*) and *neuropeptide B* (*npb*) or the calcium regulated factor scinderin like b (*scinlb*). The alpha-enriched genes also comprise the transcription factors *etv1* and *si:ch211-145o7.3* (a forkhead domain factor). The alpha-selective expression of *pnoca* and *scinlb* was confirmed by FISH, validating the RNA-seq data (Additional file 12: Figure S5).

As for the delta cells, besides the somatostatin genes [47], the RNA-seq data revealed many novel markers, which include genes coding for transcription factors such as *dlx3b*, *cdx1*, *cdx4*, and *tbx2b*, as well as genes coding for signaling factors like *bmp7b* or the kinase Map3k15. Interestingly, as demonstrated by the GO enrichment analysis (Additional file 10: Figure S3), many GPCR are specifically enriched in delta cells such as npy8br (neuropeptide Y receptor), glra4a (glycine receptor), uts2r (urotensin receptor), ptger1b & 2a (prostaglandin receptor), adra1d (adrenoreceptor), grm3 (glutamate receptor), oprd1b (opioid receptor), and gpr123, pinpointing these endocrine cells as targets of diverse external metabolic signals. We confirmed the selective expression for several genes, including the *cdx4*, *cdx1*, *lamc2*, and *map3k15* genes, detected already at 24 hpf in delta cells (Additional file 13: Figure S6). All these results validate the RNA-seq data and show that many endocrine cell subtype markers acquire their selective expression in the first embryonic endocrine cells.

Comparison of transcriptomic signatures of alpha and beta cells across species

The zebrafish RNA-seq data led to the identification of many novel markers for each endocrine cell subtypes raising the question of whether such markers display the same cell subtype-specific expression in other species. As the transcriptome of human and murine alpha and beta cells have been determined by a similar approach using highly purified FACS cell preparations [12, 48], this allowed us to perform an interspecies comparison. In order to identify all alpha and beta cell differentially expressed genes (with no restriction on endocrine enriched genes), we performed exactly the same procedure on the zebrafish, mouse, and human RNA-seq datasets, namely a comparison of the entire alpha and beta cell transcriptomes within each species using the same software and selection criteria (ratio of alpha vs. beta cells above four-fold enrichment, with adjusted $P < 0.05$). This led to the identification of 747, 1330, and 1102 alpha-enriched genes as well as 544, 381, and 465 beta-

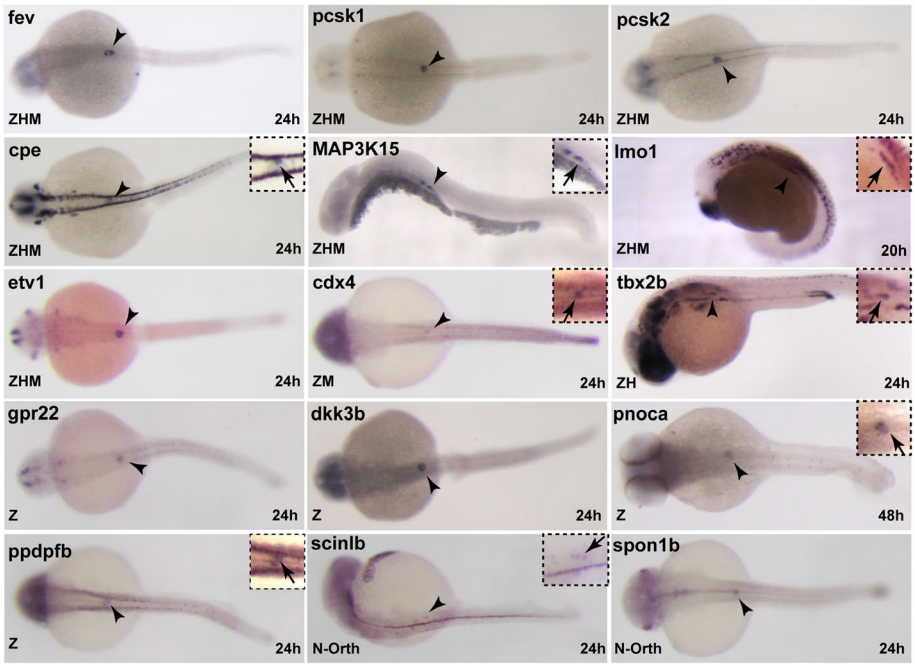


Fig. 4 Expression of genes in endocrine cells of the dorsal pancreatic bud. Whole-mount in situ hybridization on zebrafish embryos showing endocrine pancreatic expression of some new zebrafish endocrine markers (n > 15). Genes with conserved endocrine-enriched expression in zebrafish, human, and mouse (ZHM), or endocrine-enriched in zebrafish and mouse (ZH), or in zebrafish and human (HM) are indicated. Z: Gene endocrine enriched only in zebrafish. N-orth: Endocrine enriched zebrafish gene with no obvious human or mouse ortholog. Arrowheads indicate the location of the dorsal pancreatic bud containing embryonic endocrine cells and insets at the top-right display higher magnification view of the pancreatic bud

enriched genes in human, mouse, and zebrafish, respectively. Surprisingly, the interspecies comparison of these sets of genes revealed that the large majority of alpha- and beta- enriched genes are species specific (Fig. 7). Indeed, among the 465 zebrafish beta-enriched genes, only three have maintained beta cell preferential expression in human and mice, namely *insulin*, the transcription factor *pdx1*, and the glucagon receptor *gcgra* (Table 3 and Additional file 14: Table S8 for all ZH, ZM, and HM beta-enriched genes). Similarly, among the 1102 zebrafish alpha cell-enriched genes, only 20 have human and murine orthologs with alpha cell-enriched expression (Fig. 7, Table 3 and Additional file 15: Table S9). As expected, *glucagon* and *arx* are part of the conserved alpha cell genes. The *adcy2* and *adcy7* genes coding for adenylate cyclase are also alpha-enriched in the three vertebrate species supporting the important role of cAMP in alpha cells [49]. The *gc* gene (coding for vitamin D binding protein) [14] and *fev* (coding for a Ets transcription factor) are among the conserved alpha cell signatures, suggesting a specific function in this endocrine cell subtype. Taken together, all these analyses indicated that, while hundreds of genes display conserved expression in endocrine and exocrine cells, there is much less conservation for the genes differentially expressed between the endocrine cell subtypes alpha and beta. These data confirm the striking differences between species as previously reported for the human and murine alpha- and beta-enriched genes [12, 21]. Importantly, many of the conserved alpha- and beta-specific genes identified in the present study were also recently highlighted as enriched in these endocrine cell types in several single cell RNA-seq studies [19–21] (Additional file 14: Table S8 (beta cells) and Additional file 15: Table S9 (alpha cells) for comparison).

Role of the zebrafish *myt1b* and *cdx4* genes in endocrine cell differentiation

To validate our cross-species approach to identify important pancreatic regulatory factors, we selected two transcriptional regulators that were expressed at high levels in endocrine cells at the early developmental stage (27 hpf) and performed loss of function studies. Mutations of *myt1b* gene ("ZMH" conserved endocrine genes) and of the *cdx4* gene ("ZM" conserved endocrine gene) were found to affect endocrine cell differentiation. As *cdx4* is selectively expressed in delta cells, we analyzed the expression of several novel delta cell markers in the $cdx4^{tv205}$ null zebrafish mutant [50]. A previous study has shown that the zebrafish *cdx4* mutant displays defects in the anterio-posterior patterning of the endoderm, with notably a posteriorly-shifted pancreas and an increase of pancreatic beta cell number [51]; however,

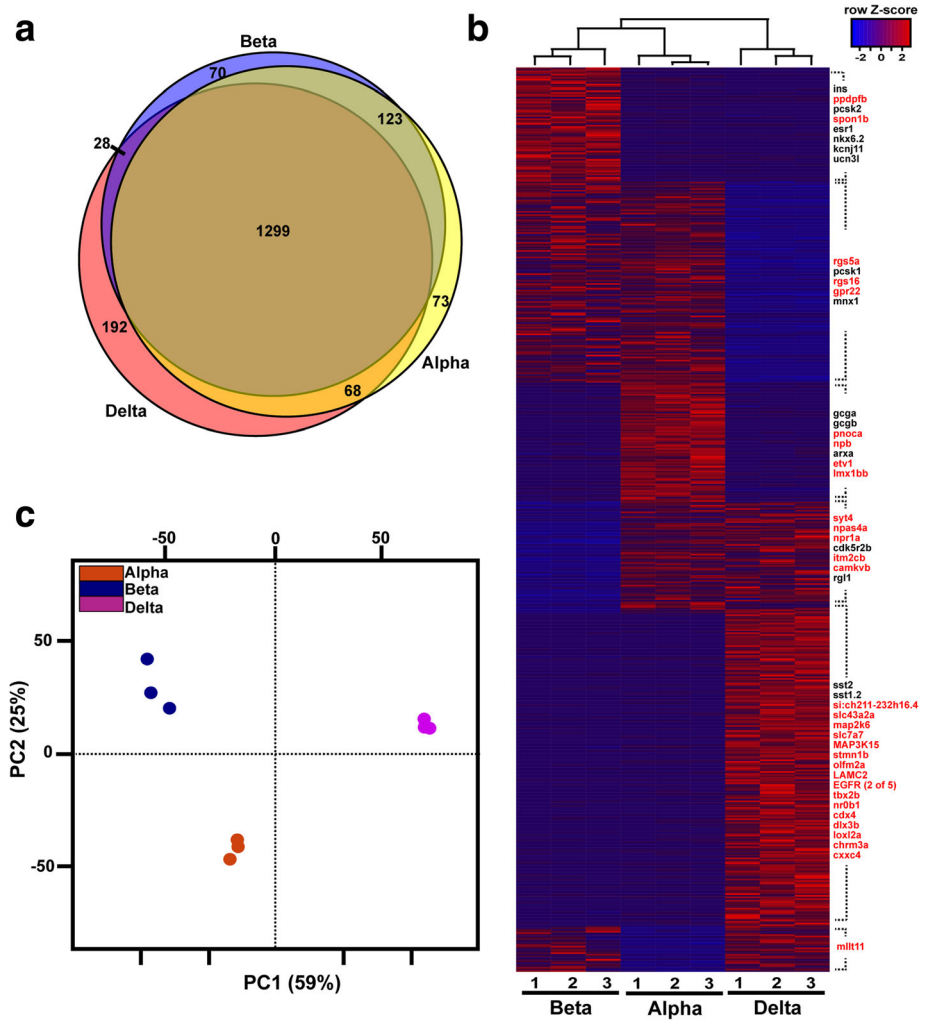


Fig. 5 Zebrafish genes differentially expressed in the endocrine cell subtypes. **a** Venn diagram displaying the number of endocrine genes with alpha-, beta- and delta-enriched expression identified with DESeq algorithm based on a cut-off ratio of four-fold and adjusted $P < 0.05$. **b** Heatmap plot showing the expression pattern of all differentially expressed genes. Genes listed at the right side of the plot are some examples of either known (black) or new (red) markers identified in this analysis. **c** Principal component analysis performed on the nine zebrafish endocrine RNAseq datasets using the 1853 endocrine-enriched genes. Compared to the PC plot of Fig. 1b performed on all annotated zebrafish genes (33,726 genes), this plot shows a tighter clustering of all replicates and better discrimination between the three endocrine cell subtypes

the selective expression of *cdx4* in delta cells was unknown as well as its function in this endocrine cell type. Analysis of the null *cdx4*^{tv205} mutant embryos revealed a complete loss of *somatostatin2* gene expression (Fig. 8a). Similarly, expression of the new delta cell markers *lamc2* and *slc7a7* was almost undetectable in the mutant embryos, in contrast to the increase of insulin expression (Fig. 8b, c). Expression of *map3k15* gene in delta pancreatic cells was also specifically abrogated by *cdx4* mutation, while expression of this gene was not affected anteriorly at the level of the pronephric glomeruli (white arrows Fig. 8d). All these data demonstrate the important role of *cdx4* for delta cell differentiation.

Our transcriptomic analyses show that embryonic pancreatic cells express both *myt1a* and *myt1b* paralogs, with *myt1b* being expressed at much higher levels. This was confirmed by ISH revealing expression of *myt1b* in the dorsal pancreatic bud of zebrafish embryos (Additional file 16: Figure S7), while the paralog *myt1a* was barely detectable (data not shown). The adult RNA-seq data indicate that the *myt1a/b* genes are expressed in alpha, beta, and delta cells. Simultaneous inactivation of the two zebrafish *myt1* genes was performed by multiplex CRISPR/Cas9 mutagenesis through injection of four guide RNA (two different guides targeting each gene, see Methods). Injected embryos revealed a significant decrease of *glucagon* expression and no significant effect on *insulin* expression (Additional file 16: Figure S7). To confirm these data obtained in F0 embryos, a line harboring a null mutation in *myt1b* (*myt1b*^{ulg029}) was raised.

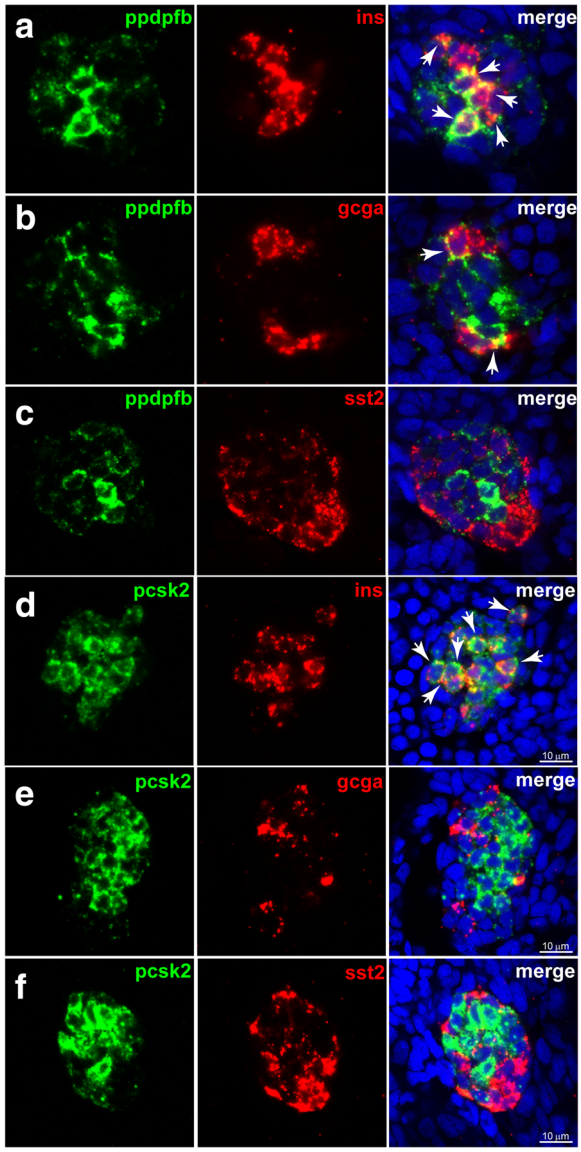


Fig. 6 Expression of *ppdpfb* and *pcsk2* genes in zebrafish beta cells. Co-labeling by fluorescent *in situ* hybridization (FISH) of *ppdpfb* and *pcsk2* with insulin, glucagon, and somatostatin. **a**–**c** *ppdpfb* is mainly expressed by beta cells (**a**, arrows) and in few alpha cells (**b**, arrows). No expression of *ppdpfb* was observed in delta cells (**c**). Beta cells specifically expressed *pcsk2* (**d**, arrows) while no expression was detected in alpha or delta cells (**e** and **f**). (Analyzed embryos > 10)

Alpha cell mass was decreased in the *myt1b*$^{ulg029-/-}$ embryos compared to the wild-type and heterozygous siblings (Fig. 9).

Discussion

In this study, we have explored the transcriptomic landscape of the major pancreatic cell types in zebrafish, thereby defining the transcriptomic signature of acinar and endocrine cells, as well as of the three major endocrine cell subtypes alpha, beta, and delta. By this analysis, we identified many novel cell type-specific markers with still unknown pancreatic function. By comparing the endocrine and exocrine transcriptomic signatures from zebrafish, human, and mouse, we could define an evolutionarily conserved signature for these two pancreatic tissues, pinpointing genes and pathways that likely represent key players in the pancreas of vertebrates. Consistent with this notion, more than half of all transcription factors in these endocrine or exocrine conserved signatures are known regulators of pancreatic cell differentiation, such as the transcription factors *neuroD*, *isl1*, *pax6*, *insm1*, *ptf1a*, and *rbpjl*, among others. *Myt1* is part of the endocrine conserved signature and we show here that inactivation of the *myt1b* gene in zebrafish leads to a decrease of alpha cell mass in 2 dpf embryos. Additionally, *cdx4*, which is selectively expressed in delta cells, is also necessary for their differentiation.

The endocrine signature conserved among the three vertebrate species reveals that several signaling pathways regulating hormone secretion and cellular homeostasis are commonly used by pancreatic cells from fish to mammals. For example, the evolutionarily conserved signature comprises the RNA binding protein *Elavl4/Hud* and the ionotropic glutamate receptor *Gria2*, shown to regulate hormone synthesis and secretion in rodents [40, 52, 53]. Similarly, the glucose response system regulating insulin and glucagon secretion in mammals seems to be also used in zebrafish as various components of this pathway are found in the conserved endocrine genes such as the ATP sensitive K^+ channels *abcc8* and *kcnj11* and the voltage-dependent Ca^{++} channel *cacna2d2*. The use of this pathway in zebrafish is further strengthened by the very high expression of the glucose transporter glut2 (*slc2a2*) and the glucokinase (*gck*) in zebrafish beta-endocrine cells. Additionally, some GPCR receptors which have been reported to control the activity of pancreatic islet cells in mice, such as the Sstr3 (somatostatin receptor) [54], CasR [55], or Celsr3 [56], are included in the conserved endocrine signature. Thus, this list of conserved endocrine-enriched genes strongly suggests that the function of several signaling pathways controlling the formation and activity of pancreatic cells has been maintained from fish to humans. This is consistent with many studies using zebrafish as a model for pancreas development and diabetes that revealed conserved physiological regulations (reviewed in [34, 35]). The conserved endocrine and exocrine pancreatic signatures also pinpoint to many novel candidate regulatory genes that deserve special attention for future studies. For instance, the transcription factors *lmo1* and *npas4a*, or the signaling factors *gpr158*, *rgs7*, and *rgs17*, all selectively expressed in pancreatic endocrine cells in zebrafish, mice, and human, more than likely play an important role in pancreatic cells.

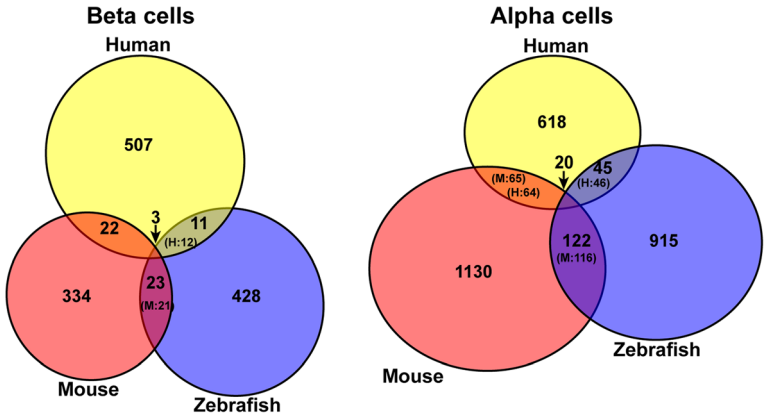


Fig. 7 Identification of genes with conserved enriched expression in alpha and beta cells. Venn diagrams showing the number of genes presenting enriched expression in alpha cells (right panel) and beta cells (left panel) and displaying this enrichment in zebrafish, mice, and/or human (shown in intersections). Due to gene duplications in some species and often in zebrafish, the number of corresponding murine (M) or human (H) orthologous genes is given in brackets in each intersection. The alpha- and beta-enriched genes were selected by DESeq2 with fold change > 4 and adjusted $P < 0.05$ using two murine alpha and beta cell preparations [12], six human alpha and beta cell preparations [48], and three zebrafish alpha and beta cell preparations (this study). The full list of conserved beta- and alpha-enriched genes is given in Additional file 14: Table S8 and Additional file 15: Tables S9

In the present study, we identified the human and murine endocrine- and exocrine-enriched genes by comparing transcriptomic data from purified islets with data obtained from whole pancreas as well as from one human pancreatic sample enriched in acinar tissue. While this approach identified almost all expected endocrine and exocrine markers, some markers may have been missed due to the low purity of tissue preparations. For example, the human *MNX1* and murine *Esr1* (*Estrogen receptor alpha*) genes were not enriched in the human or murine endocrine samples, respectively, while these two genes were reported to be important for endocrine beta cells in both species [57–60]. These misclassifications may be due to the use of RNA-seq data obtained from whole pancreas instead of highly purified exocrine cells; such potential errors will be corrected when RNA-seq data from purified ductal and acinar cells will be available. The recent single cell transcriptomic data reported for human and murine pancreas bypass the need for purified cells and will probably constitute a useful resource for doing such interspecies comparison. However, murine acinar cells have not been captured in these single-cell studies preventing the identification of acinar-enriched and endocrine-enriched genes. When such data is available, it will be interesting to perform a global analysis of all pancreatic single cell transcriptomic studies and compare it with the zebrafish pancreatic data. In the meantime, genes classified in our list of conserved "ZM" or "ZH" endocrine genes should be also considered as they indeed include *mnx1* and *esr1*, indicating that important genes fall in these two categories. For proof, *cdx4*, classified as endocrine-enriched in zebrafish and mice (i.e., "ZM"), was found here to be crucial for endocrine delta cell differentiation as many novel delta cell makers were drastically reduced in the *cdx4* mutant zebrafish embryos. Our findings warrant future analyses on the murine (and human) *Cdx4* gene to decipher its expression and function during pancreas development in mammals. Interestingly, the loss of delta cells with the concomitant increase of beta cells observed in the zebrafish *cdx4* mutant suggests that Cdx4 could act on the balance of delta versus beta cells by determining the fate of endocrine precursors.

The present study also provides a comprehensive list of new markers of the zebrafish endocrine alpha, beta, and delta cell subtypes. An overall view indicates that the alpha and beta cells are slightly more similar to each other compared to delta cells, as shown by the PC analysis (Figs. 1b and 5c). This is also supported by the Venn diagram (Fig. 5a) showing that (1) the number of genes with enriched expression in both alpha and beta cells and not in delta cells ("alpha-beta" enriched genes) is higher than those of the "alpha-delta" enriched and "beta-delta" enriched genes (123 genes vs. 28 and 68 genes, respectively), and (2) the number of delta cell markers is higher than the alpha and beta cell marker (192 vs. 70 and 73 genes, respectively). These observations are consistent with a recent study in zebrafish larva showing the higher capacity of alpha cells to trans-differentiate to beta cells compared to delta cells [61]. Whether this is also true in adult zebrafish must be verified as it has been demonstrated in mice that the trans-differentiation competence of alpha or delta cells toward beta cells is age dependent [62].

An unexpected observation of our study is the low conservation of the alpha and beta cell transcriptomic signatures between zebrafish, mice, and human. Indeed, only 20 and three genes defined the conserved alpha and

Table 3 List of genes with evolutionary conserved alpha and beta cell-enriched expression

Conserved beta cell-specific genes					
Zebrafish	Log2FC	Mouse	Log2FC	Human	Log2FC
pdx1	5.4	Pdx1	2.9	PDX1	6.4
gcgra	5.1	Gcgr	2.6	GCGR	4.2
ins	5.7	Ins2	5.4	INS	6.9
ins	5.7	Ins1	5.3	INS	6.9
Conserved alpha cell-enriched genes					
Zebrafish	Log2FC	Mouse	Log2FC	Human	Log2FC
arxa	12.1	Arx	6.7	ARX	6.0
nrxn3a	4.3	Nrxn3	3.1	NRXN3	2.5
adrb1	2.3	Adrb1	4.3	ADRB1	4.8
fap	2.4	Fap	3.0	FAP	7.1
adcy2a	6.0	Adcy2	5.7	ADCY2	3.2
grap2a	5.6	Grap2	2.3	GRAP2	2.4
ptprz1b	2.3	Ptprz1	5.0	PTPRZ1	4.6
gcgb	7.4	Gcg	5.6	GCG	7.5
gcga	5.5	Gcg	5.6	GCG	7.5
adcy7	3.5	Adcy7	3.7	ADCY7	2.2
adamts18	3.5	Adamts18	5.4	ADAMTS18	5.0
gata6	6.8	Gata6	3.0	GATA6	4.3
gc	4.9	Gc	2.5	GC	6.6
fev	4.3	Fev	5.0	FEV	5.5
tgfbr2	5.6	Tgfbr2	2.6	TGFBR2	2.3
kcnc2	7.8	Kcnc2	4.0	KCNC2	4.9
fgb	2.9	Fgb	2.5	FGB	3.0
marcksl1a	5.4	Marcksl1	2.7	MARCKSL1	2.1
mrc1b	5.5	Mrc1	2.2	MRC1L1	2.9
tgm2b	4.2	Tgm2	2.6	TGM2	2.1

Expression enrichment in alpha and beta cells is given as the Log2 fold-change in the three species

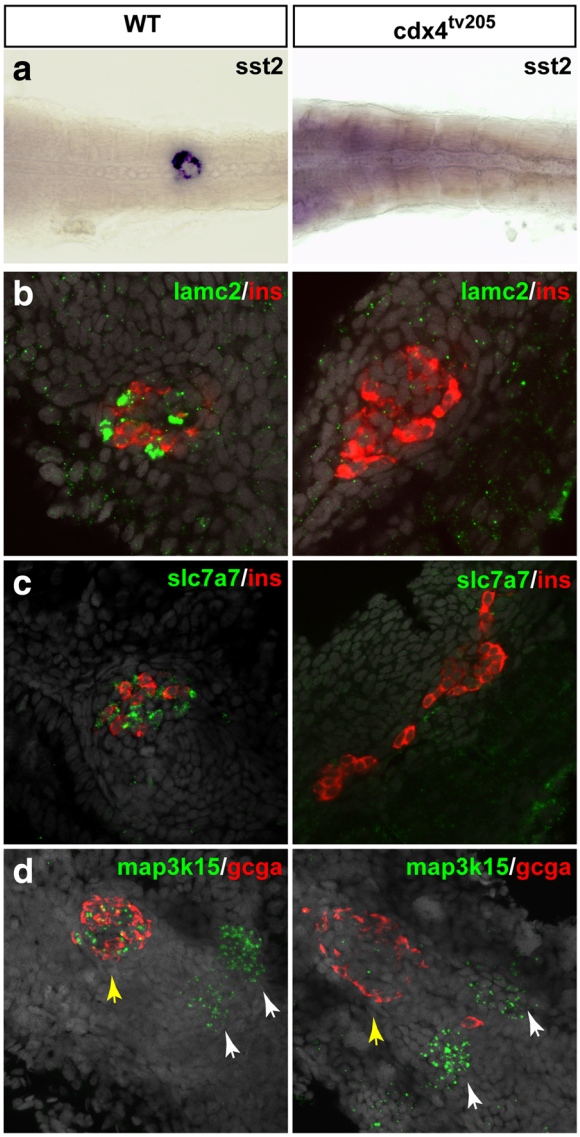


Fig. 8 Delta cell differentiation is disrupted in the cdx4 mutant embryos. Analysis of wild-type and *cdx4*$^{tv205-/-}$ mutant embryos at 48 hpf by WHISH (**a**) and FISH (**b**–**d**) using delta cell markers (*sst2*, *lamc2*, *slc7a7*, and *map3k15*) as well as *gcga* and *insulin*. *cdx4*^{tv205} mutants display a loss of *sst2* (**a**), *lamc2* (**b**), and *slc7a7* (**c**) pancreatic expression and an increase of beta cells. *Map3k15* expression is strongly reduced in the pancreatic islet (yellow arrows) while not affected in the presumptive pronephric glomeruli (white arrows) (**d**). No obvious effect is observed on *gcga* expression. Nuclei are stained with DAPI (grey staining). (Analyzed embryos > 10)

beta cell signatures, respectively. Such low conservation cannot be attributed to a low quality of cell samples or RNA-seq data because, for the three species datasets, there is a very high enrichment of alpha and beta cell markers in the corresponding libraries confirming the good purity of alpha and beta cell preparations. Further, it is noteworthy that this low conservation is not only observed between zebrafish and the two mammalian species but also between human and mouse (see Venn diagram in Fig. 8). Differences between human and mouse alpha- and beta-enriched genes have also been previously noticed [12, 21]. Several non-exclusive explanations can be proposed for this apparent low conservation of the alpha- and beta-signatures in both distant and closer species. First, as the alpha and beta cells have relatively similar transcriptomes, as shown on Fig. 1a, the identity of these two cell subtypes could rely on the action of only few regulatory genes. This hypothesis is supported by the presence of *Arx* among the few conserved alpha cell genes, which is the key determinant of alpha cell identity and its inactivation in mice is sufficient to transdifferentiate alpha to beta cells [63]. Similarly, Pdx1, one of the three conserved beta cell genes, is required to maintain beta cell identify and

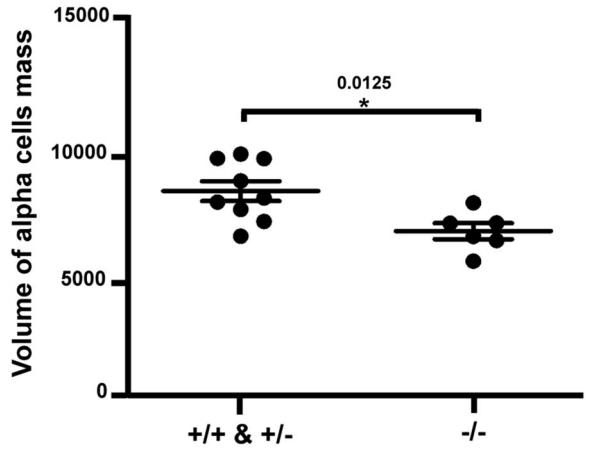


Fig. 9 Reduced glucagon expression in the *myt1b* zebrafish mutant embryos. Glucagon expression was analyzed by immunofluorescence in F2 *myt1b*^{ulg019} homozygous mutant larvae and in sibling $^{+/+}$ (WT) and $^{+/-}$ larvae at 2 dpf. The volume of the alpha cell mass was determined in each embryo by the imaging software Imaris. The graph shows the quantification measured for all myt1b$^{-/-}$ embryos and for $^{+/-}$ or $^{+/+}$ siblings, indicating a slight by statistical significant decrease of alpha cell mass (each point is the alpha cell mass measured in one embryo)

repress the alpha cell program [64]. This first hypothesis is also supported by the ability of alpha and beta cells to transdifferentiate into each other [65, 66]. A second explanation can be the relatively stringent threshold that we used to select differentially expressed genes (four-fold enrichment, adjusted $P < 0.05$). For example, the *mnx1* gene, known to be involved in beta cell differentiation [67, 68], is enriched in beta cells in the three vertebrate species but was not classified as beta-enriched in human and zebrafish due to a rather low beta versus alpha cell enrichment (three-fold in human, 1.9-fold in zebrafish versus 70-fold in mice). These differences in enrichment between human and mouse *Mnx1* is confirmed in the recent single cell RNA-seq data [18–21]. Similarly, the *prohormone convertase 1* gene *pcsk1* required for insulin hormone maturation was only 2.7-fold enriched in zebrafish beta cells versus alpha cells. It is possible that small differences in expression levels may be sufficient for some genes to give a different physiological response. A third explanation of the low cell subtype conservation could stem from functional switches occurring between homologous genes able to perform the same function. For example, *Nkx6.1* is expressed in beta cells in mice and human but not in zebrafish, where its paralog *nkx6.2* instead of *nkx6.1* is selectively expressed in mature beta cells [45]. *nkx6.2* may fulfill the function of *Nkx6.1* as Nkx6.1 and Nkx6.2 have equivalent biological activities [69]. Another example is the Zn^{++} transporter *scl30a8*, which is strongly expressed in beta cells in mammals but not in zebrafish, which instead strongly expresses the homologous Zn^{++} transporter *scl30a2*. Such functional switches may be difficult to identify as they can occur between genes of the same superfamily but belonging to different subclasses. For example, we have shown that, in zebrafish, the role of the bHLH Neurog3 as pancreatic endocrine cell fate determinant is fulfilled by two other bHLH factors Ascl1b and Neurod1 [70]. A functional switch is also observed for the specification of the secretory cell of the intestine, played by Atoh1 in mice and by Ascl1a in zebrafish [71]. The capacity of homologous factors to fulfill the same function is also supported by the observations that many null mutations can be compensated by a homologous gene [72–74]. Such compensations seem to occur for the murine *Myt1* gene as the *Myt1* KO mice only display very mild pancreatic defects with no decrease of endocrine cells [75], while more drastic defects are observed by expression of a dominant negative Myt1-Eng protein [1, 76]. Furthermore, adult murine endocrine pancreatic cells express the homologous *Myt1l* and *Myt3/St18* genes [77] and their expression increases in the *Myt1* KO mice [75]. In zebrafish, *myt1a*, *myt1l*, and *myt3/st18* are expressed at much lower level compared to *myt1b*; this probably explains the decrease in alpha cell mass observed in the single *myt1b* mutant. Further experiments will be required to determine whether some compensation occurs through *myt1a* and if the defects are more drastic in the double *myt1a/b* mutant.

The stronger expression of glucagon receptor in the beta cells of the three vertebrate species highlights the importance of paracrine regulations between alpha and beta cells and indicates a role of glucagon on beta cell physiology, as previously reported in rodents [78] and in zebrafish for beta cell regeneration [61]. Interestingly, our transcriptomic analyses highlight *fev* as an endocrine conserved gene as well as an alpha cell conserved gene, suggesting a role of this transcription factor in alpha cells. Glucose tolerance test in *Fev* knockout mice previously revealed a slower response apparently due to a reduction in insulin production while the level of most crucial transcription factors of beta cells were unchanged [79]. As our transcriptomic analysis indicates that *fev* is expressed more than 20-fold higher in alpha cells versus beta cells in zebrafish, mice, and human, this argues for an important function in alpha cells and warrants further phenotypic analyses.

Conclusions

The present transcriptomic analysis of the distinct zebrafish pancreatic cell types identifies novel pancreatic regulatory genes and thereby constitutes a valuable resource for future studies of the pancreas not only in zebrafish but also in mammals giving interesting clues on genes and signaling pathways active in these cells.

The comparison of the present zebrafish pancreatic RNA-seq data with the recent single cell pancreatic transcriptomic data or with future data of pure pancreatic cells obtained not only from human and mouse but also from other species will help to better define gene regulatory networks controlling pancreas ontogeny and physiology in vertebrates. This will be useful for studies aimed at understanding dysfunction of pancreatic endocrine cells in human diseases like diabetes and to design novel drugs or therapies.

Methods

Generation of the Tg*(Insulin:GFP)*$^{ulg021\ Tg}$ line

The *(Insulin:GFP)* transgene was generated by first cloning a 897 pb PCR fragment, amplified with O97 and O98 primers (Additional file 17: Table S10), that includes 745 bp of the insulin promoter, the exon 1 and intron 1 and 7 bp of exon 2 just upstream of the ATG of the insulin ORF, into the gateway vector pCR8/GW/TOPO to produce pMV90-G2a plasmid. Second, a triple LR recombination using p5E-MCS, pMV90-G2a, and p3E-EGFP-PA inserted into pDestTol2pA2, provided by the tol2kit [80], generated the transgene *Tg(insulin:GFP)* that has been introduced into AB embryos by coinjection with the Tol2 transposase to generate the *Tg(insulin:GFP)*$^{ulg021\ Tg}$ line.

Preparation of zebrafish pancreatic cells by FACS

The endocrine cells were prepared using the zebrafish transgenic lines *Tg(Insulin:GFP)*^{ulg021Tg}, *Tg(gcga:GFP/ins:mCherry)ia1*, and *Tg(st22:GFP)* to isolate beta, alpha, and delta cells, respectively [28, 81]. Acinar and ductal cells were isolated by using *Tg(ptf1a:GFP)* [27] and *Tg(nkx6.1:GFP)*^{ulg004Tg} [29], respectively. The endocrine tissue was dissected under an epifluorescence stereomicroscope before dissociation by enzymatic treatment for 30 min at 28 °C with 1× Tryple Select (Life Techologies), 40 μg/mL proteinase K (Roche), and 10 μg/mL collagenase IV (Life Technologies), combined with mechanical disruption by pipetting every 5 minutes. For acinar cell preparations, pancreata were digested with a mix of collagenase IV (500 μg/mL), collagenase P (350 μg/mL), and dispase II (1 mg/mL) for 15 min at 28° with mechanical disruption every 5 minutes. After dissociation, cells were washed twice with 1× PBS containing 1% BSA and pelleted at 4 °C for 5 min at 300 *g* and immediately sorted by FACS. Cells were selected based on GFP expression using FACS Aria II by two consecutive sorting steps: the first sorting was done in the "yield" mode and the second in the "purity" mode. Purity was estimated by FACS Aria II after cell sorting (more than 99% purity of GPF^+ cells) and by fluorescence microscopy (from 95–99% purity). Each replicate sample was prepared from four adult zebrafish. Approximately 10,000–20,000 endocrine cells, 20,000–40,000 acinar cells, and 2000–10,000 ductal cells were obtained after FACS and used for library preparations.

RNA extraction, cDNA amplification, library preparation, and sequencing

Total RNA was extracted from FACS sorted cells using the RNeasy plus micro kit (Qiagene). RNA from endocrine and acinar cells was eluted in 10 μL with a concentration of 100–400 pg/μL. RNA integrity was assessed by a capillary electrophoresis using Agilent RNA 6000 pico chip (Agilent technologies), the RIN value for each sample was from 8 to 10. The Smarter Ultra low RNA input kit (clontech) [82] was used to for the synthesis and amplification of cDNA synthesis using up to 10 ng of total RNA following the manufacturer's instructions and performing no more than 12 cycles of PCR in order to minimize amplification biases. The quality of cDNA was verified by 2100 High Sensitivity DNA assay (Agilent technologies). Truseq DNA Illumina libraries were prepared and sequenced to obtain approximately 90 million reads (100 bp paired-end reads) per library using the Hiseq 2000 Illumina sequencer.

RNA-seq data analysis

Sequences were trimmed in order to remove adaptors and low quality bases. Trimmed reads were mapped in to the genome (Zv9, Ensembl genome version 75) using Tophat v.2.0.9 [83]. Tophat's options were set according to the library features (-r 220 –mate-std-dev 82 –segment-length 18) and the option –min-intron-length was set up to 30 nucleotides according to the intron length described by Moss et al. [84]. For the mouse and human datasets, raw data were downloaded from the public databases: human pancreas (four samples from Fagerberg et al. [37] and one acinar cell-enriched sample from Morán et al. [9]: the E-MTAB-1294 dataset), human islets (three samples from Nica et al. [46] and four samples from Morán et al. [9]), murine pancreas (two samples from Holmstrom et al. [36]), and murine islets (three samples from Morán et al. [9]). The alpha and beta cell RNA-seq datasets were obtained from Benner et al. [12] for mouse (two samples of each) and from Blodgett et al. [48] for human (six samples each) [48]. Reads were mapped into the mouse GRCm38 and the human GRCh37 genomes (Ensembl genome version 75) using default options. Gene expression was measured from the mapped reads by using HT-seq-count [85]. PCA, using princomp function of R, was calculated for the whole dataset using the variance stabilization transformation values obtained by DESeq R package from the gene expression values [86]. For differential expression analysis, we used the R package DESeq2 [87]. DESeq2 employs shrinkage estimation for dispersions and fold change; it uses Wald test for significance with posterior adjustment of *P* values using the procedure

of Benjamini and Hochberg (giving adjusted P values). Genes differentially expressed were selected with an adjusted $P < 0.05$ and a fold change > 4. For the comparison of the acinar, ductal and endocrine cell data, in-silico endocrine datasets were simulated by combining alpha, beta and delta RNA-seq data. This approach was taken to decrease the number of pairwise comparisons and could be used since the sequencing deepness of each library was in a similar range (below of a factor 2). Endocrine dataset #1 was obtained by combining the mapped reads of alpha1 (47×10^6 reads), beta1 (48×10^6 reads), and delta1 (31×10^6 reads) datasets; endocrine dataset #2 is a mix of alpha2 (46×10^6 reads), beta2 (29×10^6 reads), and delta2 (30 10^6 reads); endocrine dataset #3 is the combination of alpha3 (59×10^6 reads), beta3 (62×10^6 reads), and delta3 (59×10^6 reads) datasets. The RNA-seq raw data have been deposited on ENA under the accession number PRJEB10140.

Comparison of the human, murine, and zebrafish pancreatic transcriptomes

The predicted orthologs among zebrafish, mouse and human were obtained from Ensembl [88]. The information from the three species was retrieved using Biomart tool [89]. Two different orthology tables were generated (tables available upon request). The first table contains all zebrafish, murine and human genes with 1-1-1 orthology relationship. Orthology table two comprises all 1-1-1 orthologs as well as the genes presenting 1-many-many orthology relationships and thus includes all duplicated genes (paralogs), which are notably found in the zebrafish genome. The interspecies PC analysis was performed using only the genes with a 1-1-1 orthology relationship. The genes presenting an endocrine-enriched or an exocrine-enriched expression were identified in mouse and in human by using DESeq2 software selecting genes with at least four-fold higher expression in islet dataset or in whole pancreas dataset (adjusted $P < 0.05$). The interspecies comparison of endocrine- and exocrine-enriched genes was performed using the orthology table two.

GO enrichment analysis

Tissue- and cell type-enriched genes were converted and uploaded to DAVID bioinformatics resource [90]. Using the Functional annotation tool, we run a GO enrichment analysis for endocrine and acinar enriched genes as well as for the alpha-, beta- and delta-enriched genes. Enriched processes were identified by GOTerm_BP_FAT with a $P < 0.1$.

In situ hybridization

Antisense RNA probe for the different genes were prepared as described by Thisse et al. [91], except for *fev* [92]. Briefly, primers were designed to amplify a part of the transcript that is used as a template to synthesize the probe. The reverse primer at the 5' end contains the minimal promoter sequence for T3 RNA polymerase (5'-AATTAACCCTCACTAAAGGGAG-3'), templates were amplified by RT-PCR using the set of primers shown in Additional file 17: Table S10. Whole mount in situ hybridization and fluorescent in situ hybridization (WISH and FISH) were performed as described by Mavropoulos et al. [93], applying some modification to this protocol. Briefly, larvae of 3 dpf or older were incubated during 20 minutes in methanol and 3% H_2O_2 at room temperature, prior to dehydration. Antisense probe hybridization was performed using 10–50 ng of DIG- and DNP-probes in hybridization buffer containing 5% dextran sulfate (MW: 500,000) at 65 °C overnight. Antibodies were pre-absorbed on homogenized larvae (mix of different developmental stages) for 2 h at room temperature and then diluted to 1/3000 DIG-AP, 1/1500 DIG-HRP, and 1/800 DNP-HRP (PerkinElmer).

Inactivation of *myt1a* and *myt1b* genes by multiplex CRISPR/cas9 mutagenesis

Mutations in the *myt1a* and *myt1b* genes were generated by multiplex CRISPR/Cas9 technology essentially as described previously [94, 95]. The nls-zCas9-nls mRNA was synthesized by transcription of the plasmid pT3TS-nCas9n (Addgene). CRISPR guide RNAs were selected using CRISPR design and chopchop software to target the beginning of Myt1a and Myt1b coding regions (guides 1) and the regions coding for the first zinc finger domain (guides 2). The following target sites were used: GCCAAGACGCAGATGATAAGCGG and GATGGTT TAGGCCATGTCAGTGG for *myt1a*, and GTCTGA GGGAGGGCCGGCAGCGG and TGCCATTGCATCC TGGAGTGGGG for *myt1b* (PAM motifs are underlined). The DNA templates were prepared by annealing and filling two oligonucleotides containing the T7 promoter sequence and the target sequences as previously described [95]. After synthesis and purification of gRNA, fertilized zebrafish eggs were injected with approximately 1 nL of a solution containing 50 ng of the four gRNA and 300 ng of nls-zCas9-nls mRNA. The efficiency of mutagenesis was verified by genotyping using Heteroduplex Migration Assays after amplification of targeted genomic sequences. A few injected embryos were fixed in PFA at 48 hpf for phenotypic analysis and the others were raised until adulthood. Founder fish transmitting a germline mutation in myt1b were outcrossed with wild type fish; F1 fish harboring a 5 bp insertion in *myt1b* (*myt1b*^{ulg039} allele) causing a frameshift in the coding sequences were incrossed to generate *myt1b*$^{-/-}$ embryos and $^{+/-}$, $^{+/+}$ siblings, which were analyzed by immunohistochemistry.

Immunohistochemistry

Expression of insulin and glucagon was analyzed by immunofluorescence on whole-mount zebrafish embryos. After overnight fixation in 2% PFA at 4 °C, 1 hour incubation in PBS 1% Triton X-100 for 2 hours at room temperature in blocking solution (4% BSA, 10% DMSO, 0.3% Triton X-100 in PBS), embryos were incubated with the primary antibodies anti-Mouse Glucagon 1/200 (Sigma, G2654) and Anti-Guinea Pig Insulin 1/300 (MP, 64714). After washing, embryos were incubated with the secondary antibodies Anti-Mouse Alexa 568 or 633 (Invitrogen, A-11004 and A-21052) and Anti-Guinea Pig Alexa 568 or 633 (Invitrogen, A-11075 and A-21105) diluted 1/300. After washing and mounting, embryos were scanned with a Leica SP5 confocal microscope and images were analyzed using Imaris 7.2.3 software. Alpha and beta cell mass was measured using Imaris based on 3D reconstitution of Fluorescence signal obtained by immunofluorescence; the same confocal and Imaris parameter settings were used for wild-type and mutant embryos. To compare the number of alpha cells and beta cells in wild-type and mutants, glucagon$^+$ and insulin$^+$ cells were counted in every 6-μm optical section throughout the whole principal islet. Expression of GFP and mCherry in the pancreas of Tg(gcga:GFP);(ins:NTR-mCherry) adult transgenic fish as well as of Tg(sst2:GFP) adult fish was analyzed by immunofluorescence on cryosections (Additional file 1: Figure S1). The antibodies used are the same than described above and also include the anti-Rabbit Somatostatin 1/300 (Dako, a0564) and Anti-Rabbit Alexa 568 or 633 1/300 (Invitrogen, A-11011 and A-21070). Expression of GFP in beta cells was also verified for the Tg(insulin:GFP) larvae (Additional file 1: Figure S1).

Additional files

Additional file 1: Figure S1. Expression of the transgenes *Tg(ins:GFP)*, *Tg(gcga:GFP)*, *Tg(sst2:GFP)*, and *Tg(ins:NTR:mCherry)* in the endocrine pancreatic cell types. (A–C) Whole mount immunostaining of 3 dpf transgenic larvae *Tg(ins:GFP)* demonstrating the selective expression of GFP in beta cells. (D–I) Immunostaining of pancreas sections from adult transgenic fish *Tg(gcga:GFP/ins:NTR:mCherry)*. Glucagon$^+$ cells (see arrows in D–F) express high level of GFP and are not labelled by mCherry, while many insulin expressing cells are labelled by mCherry and by GFP (at slightly lower levels) (see arrowheads in G–I). These data reveal a leaking expression of the gcga:GFP transgene in beta cells. (J–L) ISH performed on pancreas section of adult Tg(sst2:gfp) with sst2 probe followed by immunofluorescence using the GFP antibody; the expression of endogenous sst2 gene co-localize with GFP staining confirming the specific expression of the transgene in delta cells [28]. (TIF 8638 kb)

Additional file 2: Table S1. Gene expression levels in alpha, beta, and delta cell subtypes. The expression is given in Normalized counts for the three replicates of alpha, beta, and delta cell libraries. (XLSX 4416 kb)

Additional file 3: Table S2. Gene expression levels in endocrine, acinar, and ductal cells. The expression is given in Normalized counts for the four replicates of acinar cells, the three replicates of ductal and endocrine cells. The three endocrine datasets were obtained by combining the reads obtained with alpha, beta and delta cell libraries as described in Methods (e.g., endocrine1 data is a mix of alpha1, beta1 and delta1; endocrine2 data is a mix of alpha2, beta2 and delta2) (XLSX 4433 kb)

Additional file 4: Table S3. Classification of genes according to their enriched expression in endocrine, acinar, or ductal cells. Expression levels were compared for each gene between endocrine, acinar, and ductal cells (pair-wise comparison with DEseq2). Classification was performed using the cut-off values of fold change > 4 and adjusted $P < 0.05$. The enrichment is given for each gene as Log2 of fold change and with the adjusted *P* value. (XLSX 2192 kb)

Additional file 5: Table S4. Genes presenting conserved endocrine expression. The ZMH excel page shows the genes with endocrine-enriched expression in zebrafish, mice and human. The ZM, ZH and HM excel pages show genes with endocrine-enriched expression in two species (ZM: zebrafish and mice; ZH: zebrafish and human; HM: human and mice). The tables give the expression mean in endocrine cells (Normalized counts) and the enrichment (FC: fold change) in the three species. (XLSX 123 kb)

Additional file 6: Figure S2. Identification of genes with evolutionary conserved and enriched expression in pancreatic exocrine cells. Venn diagram showing the number of exocrine-enriched genes found only in zebrafish, mouse or human, and those displaying conserved endocrine-enrichment in two species or in the three species (shown in intersections). Due to gene duplications in some species and often in zebrafish, the number of corresponding murine (M) or human (H) orthologous genes is given in brackets in each intersection. The full list of conserved exocrine-enriched genes is given in Additional file 7: Table S5. (PDF 4 kb)

Additional file 7: Table S5. Genes presenting conserved exocrine expression. The ZMH excel page shows the genes with exocrine-enriched expression in zebrafish, mice and human. The ZM, ZH and HM excel pages show genes with endocrine-enriched expression in two species (ZM: zebrafish and mice; ZH: zebrafish and human; HM: human and mice). The tables give the expression mean in endocrine cells (Normalized counts) and the enrichment (FC: fold change) in the three species. (XLSX 101 kb)

Additional file 8: Table S6. List of endocrine-enriched genes with an expression level above 100 counts in embryonic endocrine cells at 27 hpf. List of 911 genes with endocrine-enriched expression in adult and showing an expression level above 100 Normalized counts. Gene expression is provided for the three replicates (selection of pancreatic cells at 27 hpf from Tg(pax6:GFP)) as well as the mean expression level. (XLSX 99 kb)

Additional file 9: Table S7. Classification of endocrine-enriched genes according to their expression in alpha, beta, and delta endocrine cells. Gene expression ratio was calculated for all pair-wise comparison (alpha versus beta: A&B; beta versus delta: B&D and alpha versus delta: A&D). The enrichment is given as Log2 fold change with the adjust value. Classification was done using the cut-off threshold of Log2 fold change > 2 and adjusted $P < 0.05$. (XLSX 265 kb)

Additional file 10: Figure S3. Gene ontology (GO) enrichment analysis for endocrine cell subtypes. Left. Bar plot displaying the number of genes constituting enriched GO terms. *P* values are denoted on the bars. Right. Fold of change (in Log2) of genes constituting the most enriched GO terms, (A) GO enrichment for the 70 beta-enriched genes. (B) GO enrichment for the 73 alpha-enriched genes. (C) GO enrichment for the 192 delta-enriched genes. (B-A: pairwise beta versus alpha; B-D: pairwise beta versus delta). (TIF 13089 kb)

Additional file 11: Figure S4. Expression of *spon1b* gene in beta pancreatic cells of zebrafish embryos. Co-labeling by FISH of *spon1b* with insulin (*ins*) (A, arrows show colocalization), while no expression was detected in delta cells (B) (n > 10). *ins*: insulin, *sst2*: somatostatin 2, N-orth: Endocrine enriched zebrafish gene with no described ortholog in human and/or mouse. (TIF 5116 kb)

Additional file 12: Figure S5. Expression of *pnoca* and *scinlb* genes in alpha pancreatic cells of zebrafish embryos. Co-labeling by in situ hybridization at 30 hpf of new discovered genes with cell type-specific markers for alpha and beta cells (n > 10). A and B. *pnoca* is expressed in alpha cells (A, arrows) while no expression was detected in beta cells (B). *scinlb* was detected specifically in alpha cells (C, arrows) but not detected in beta cells (D). gcga: glucagon a, ins: insulin, Z: no endocrine gene with

no conserved expression, N-orth: Endocrine-enriched zebrafish gene with no described ortholog in human and/or mouse. (TIF 8027 kb)

Additional file 13: Figure S6. Validation of the selective expression of some genes in zebrafish delta cells. *cdx1b*, *cdx4*, *lamc2*, and *map3k15* are specifically expressed in delta cells at 24 hpf (A, C, E, G, arrows show colocalization with delta cell-specific markers, somatostatin 2; n > 10). No expression was detected for none of the genes in beta cells (B, D, F, H). *ins*: *insulin*, *sst2: somatostatin 2*, ZHM: Gene expression conserved in zebrafish, human and mouse, ZM: Gene expression conserved in zebrafish and mouse, Z: Endocrine gene with no conserved expression, N-orth: Endocrine-enriched zebrafish gene with no described ortholog in human and/or mouse. (TIF 12250 kb)

Additional file 14: Table S8. Conserved beta cell markers. The ZMH excel page shows the genes with beta cell-enriched expression in zebrafish, mice and human. The ZM, ZH, and HM excel pages show genes with beta cell-enriched expression in two species (ZM: zebrafish and mice; ZH: zebrafish and human; HM: human and mice). Beta cell versus alpha cell expression ratio is given as fold change (FC) for each species. The ZMH, ZH, and HM genes also detected as enriched in specific pancreatic cell types in the human single-cell transcriptomic studies [19–21] are also noted in the columns on the right part. (XLSX 18 kb)

Additional file 15: Table S9. Conserved alpha cell markers. The ZMH excel page shows the genes with alpha cell enriched expression in Zebrafish, Mice and Human. The ZM, ZH and HM excel pages show genes with alpha cell-enriched expression in two species (ZM: Zebrafish and Mice; ZH: Zebrafish and Human; HM: Human and Mice). Alpha cell versus beta cell expression ratio is given as fold change (FC) for the three species. The ZMH, ZH and HM genes also detected as enriched in specific pancreatic cell types in the human single cell transcriptomic studies [19–21] are also noted in the columns on the right part. (XLSX 35 kb)

Additional file 16: Figure S7. Expression and function of myt1b in zebrafish pancreas. A: Whole-mount ISH showing expression pattern of *myt1b* in zebrafish embryos at 27 hpf. High expression is detected in the dorsal pancreatic bud (indicated by the arrow) and in the central nervous system. B: Quantification of glucagon and insulin expression at 48 hpf in wild-type (non-injected) embryos and in embryos injected with the 4 CRISPR *myt1a/b* guide RNA and Cas9. Graph B shows the volume of all $gcga^+$ cells and all ins^+ cells measured in each embryo by the imaging software Imaris (see Methods) (each point is the volume measured in one embryo). This quantification indicates a statistically significant reduction of the volume of alpha cell mass while beta cell mass is not drastically affected in the injected (F0) embryos (results of one experiment). (TIF 2416 kb)

Additional file 17: Table S10. Primers used for RNA probe synthesis. List of primers used for the synthesis of the RNA probes for in situ hybridization, antisense primers contain the T3 minimal promoter sequence. (XLSX 10 kb)

Abbreviations

dpf: Days post fertilization; FISH: Fluorescent in situ hybridization; gcga: Glucagon a; GPCRs: G-couple protein receptors; hpf: Hours post fertilization; ins: Insulin; PCA: Principal component analysis; sst2: Somatostatin 2; WISH: Whole-mount in situ hybridization; ZH: Genes with conserved expression pattern in zebrafish and human; ZM: Genes with conserved expression pattern in zebrafish and mice; ZMH: Genes with conserved expression pattern in zebrafish, mice and human

Acknowledgments

We are very grateful to Francesco Argenton for the transgenic line Tg(*gcga:GFP*), to Zhen Li and Zhiyuan Gong for the transgenic line Tg(*st22:GFP*), to Steven Leach for the transgenic line Tg(ptf1a:GFP), and to Leonard Zon for the *kgg/cdx4* mutant line. We thank the following GIGA technical platforms: GIGA-Zebrafish (H. Pendeville), GIGA-Cell Imaging and Flow Cytometry platform (S. Ormenese and S. Raafat), GIGA-Genotranscriptomic (B. Hennuy, W. Coppieters and L. Karim), and GIGA-Immunohistochemistry (C. Humblet and E. Dortu). ET-S was supported by WBI, Becas Chile and Leon Fredericq fund, AL by FRIA, KP by WBI, MLV, IM, and BP are Chercheur qualifié FNRS. This work was funded by the FNRS-FRS, the Belgian State's "Interuniversity Attraction Poles" Program (SSTC, PAI) and the "Fonds Speciaux" from ULg.

Availability of data and materials

The RNA-seq raw data have been deposited on ENA (http://www.ebi.ac.uk/ena) under the accession number PRJEB10140. The gene expression values are given in Additional file 2: Tables S1 and Additional file 3: Table S2.

Authors' contributions

ET-S, MLV, IM, and BP designed the experiments and ET-S, AL, AB, KP, DB, and IM performed the experiments. ET-S, MLV, IM, and BP wrote the manuscript. All authors read and approved the article.

Competing interests

The authors declare that they have no competing interests.

Ethics approval

All animal work has been conducted according to national guidelines and all animal experiments described herein were approved by the ethical committee of the University of Liège (protocol numbers 1328).

Publisher's Note

Received: 3 November 2016 Accepted: 1 March 2017
Published online: 21 March 2017

References

1. Gu G, Wells JM, Dombkowski D, Preffer F, Aronow B, Melton DA. Global expression analysis of gene regulatory pathways during endocrine pancreatic development. Development. 2004;131:165–79.
2. Gunton JE, Kulkarni RN, Yim S, Okada T, Hawthorne WJ, Tseng Y-H, Roberson RS, Ricordi C, O'Connell PJ, Gonzalez FJ, Kahn CR, O'Connell PJ, Gonzalez FJ, Kahn CR. Loss of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. Cell. 2005;122:337–49.
3. Dorrell C, Schug J, Lin CF, Canaday PS, Fox AJ, Smirnova O, Bonnah R, Streeter PR, Stoeckert CJ, Kaestner KH, Grompe M. Transcriptomes of the major human pancreatic cell types. Diabetologia. 2011;54:2832–44.
4. Martens GA, Jiang L, Hellemans KH, Stangé G, Heimberg H, Nielsen FC, Sand O, van Helden J, Gorus FK, Pipeleers DG. Clusters of conserved beta cell marker genes for assessment of beta cell phenotype. PLoS One. 2011;6:e24134.
5. Blum B, Hrvatin SSŠ, Schuetz C, Bonal C, Rezania A, Melton DA. Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. Nat Biotechnol. 2012;30:261–4.
6. Eizirik DL, Sammeth M, Bouckenooghe T, Bottu G, Sisino G, Igoillo-Esteve M, Ortis F, Santin I, Colli ML, Barthson J, Bouwens L, Hughes L, Gregory L, Lunter G, Marselli L, Marchetti P, McCarthy MI, Cnop M. The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. PLoS Genet. 2012;8:e1002552.
7. Ku GM, Kim H, Vaughn IW, Hangauer MJ, Myung Oh C, German MS, McManus MT. Research resource: RNA-Seq reveals unique features of the pancreatic β-cell transcriptome. Mol Endocrinol. 2012;26:1783–92.
8. Benitez CM, Qu K, Sugiyama T, Pauerstein PT, Liu Y, Tsai J, Gu X, Ghodasara A, Arda HE, Zhang J, Dekker JD, Tucker HO, Chang HY, Kim SK. An integrated cell purification and genomics strategy reveals multiple regulators of pancreas development. PLoS Genet. 2014;10(10):e1004645.
9. Moran I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, Arnes L, Nakic N, Garcia-Hurtado J, Rodriguez-Segui S, Pasquali L, Sauty-Colace C, Beucher A, Scharfmann R, Van Arensbergen J, Johnson PR, Berry A, Lee C, Harkins T, Gmyr V, Pattou F, Kerr-Conte J, Piemonti L, Berney T, Hanley N, Gloyn AL, Sussel L, Langman L, Brayman KL, Sander M, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab. 2012;16:435–48.
10. Bramswig NC, Everett LJ, Schug J, Dorrell C, Liu C, Luo Y, Streeter PR, Naji A, Grompe M, Kaestner KH. Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. J Clin Invest. 2013;123:1275–84.
11. Nica AC, Ongen H, Irminger J. Cell-type, allelic and genetic signatures in the human pancreatic beta cell transcriptome. Genome Res. 2013;23(9):1554–62.

12. Benner C, van der Meulen T, Cacéres E, Tigyi K, Donaldson CJ, Huising MO, Caceres E, Tigyi K, Donaldson CJ, Huising MO. The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. BMC Genomics. 2014;15:620.
13. Li J, Klughammer J, Farlik M, Penz T, Spittler A, Barbieux C, Berishvili E, Bock C, Kubicek S. Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell types. EMBO Rep. 2016;17:178–87.
14. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Mol Metab. 2016;5:233–44.
15. DiGruccio MR, Mawla AM, Donaldson CJ, Noguchi GM, Vaughan J, Cowing-Zitron C, van der Meulen T, Huising MO. Comprehensive alpha, beta and delta cell transcriptomes reveal that ghrelin selectively activates delta cells and promotes somatostatin release from pancreatic islets. Mol Metab. 2016;5:449–58.
16. Wang YJ, Schug J, Won K-J, Liu C, Naji A, Avrahami D, Golson ML, Kaestner KH. Single cell transcriptomics of the human endocrine pancreas. Diabetes. 2016;65(10):3028–38.
17. Adriaenssens AE, Svendsen B, Lam BYH, Yeo GSH, Holst JJ, Reimann F, Gribble FM. Transcriptomic profiling of pancreatic alpha, beta and delta cell populations identifies delta cells as a principal target for ghrelin in mouse islets. Diabetologia. 2016;59:2156–65.
18. Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, Ryu JH, Wagner BK, Shen-Orr SS, Klein AM, Melton DA, Yanai I. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst. 2016;3(4):346–60. e4.
19. Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, van Gurp L, Engelse MA, Carlotti F, de Koning EJP, van Oudenaarden A. A single-cell transcriptome atlas of the human pancreas. Cell Syst. 2016;3(4):385–94. e3.
20. Segerstolpe Å, Palasantza A, Eliasson P, Andersson E-M, Andréasson A-C, Sun X, Picelli S, Sabirsh A, Clausen M, Bjursell MK, Smith DM, Kasper M, Ämmälä C, Sandberg R. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. Cell Metab. 2016;24:593–607.
21. Xin Y, Kim J, Okamoto H, Ni M, Wei Y, Adler C, Murphy AJ, Yancopoulos GD, Lin C, Gromada J. RNA sequencing of single human islet cells reveals type 2 diabetes genes. Cell Metab. 2016;24:608–15.
22. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, Albert FW, Zeller U, Khaitovich P, Grützner F, Bergmann S, Nielsen R, Pääbo S, Kaessmann H. The evolution of gene expression levels in mammalian organs. Nature. 2011;478:343–8.
23. Rebeiz M, Patel NH, Hinman VF. Unraveling the tangled skein: the evolution of transcriptional regulatory networks in development. Annu Rev Genomics Hum Genet. 2015;16:103–31.
24. Necsulea A, Kaessmann H. Evolutionary dynamics of coding and non-coding transcriptomes. Nat Rev Genet. 2014;15:734–48.
25. Pishesha N, Thiru P, Shi J, Eng JC, Sankaran VG, Lodish HF. Transcriptional divergence and conservation of human and mouse erythropoiesis. Proc Natl Acad Sci U S A. 2014;111:4103–8.
26. Shay T, Jojic V, Zuk O, Rothamel K, Puyraimond-Zemmour D, Feng T, Wakamatsu E, Benoist C, Koller D, Regev A, ImmGen Consortium. Conservation and divergence in the transcriptional programs of the human and mouse immune systems. Proc Natl Acad Sci U S A. 2013;110:2946–51.
27. Godinho L, Mumm JS, Williams PR, Schroeter EH, Koerber A, Park SW, Leach SD, Wong ROL. Targeting of amacrine cell neurites to appropriate synaptic laminae in the developing zebrafish retina. Development. 2005;132:5069–79.
28. Li Z, Wen C, Peng J, Korzh V, Gong Z. Generation of living color transgenic zebrafish to trace somatostatin-expressing cells and endocrine pancreas organization. Differentiation. 2009;77:128–34.
29. Ghaye AP, Bergemann D, Tarifeño-Saldivia E, Flasse LC, Von Berg V, Peers B, Voz ML, Manfroid I. Progenitor potential of nkx6.1-expressing cells throughout zebrafish life and during beta cell regeneration. BMC Biol. 2015;13:70.
30. Murtaugh LC. Pancreas and beta-cell development: from the actual to the possible. Development. 2007;134:427–38.
31. Ku GM, Pappalardo Z, Luo CC, German MS, McManus MT. An siRNA screen in pancreatic beta cells reveals a role for Gpr27 in insulin production. PLoS Genet. 2012;8(1):e1002449.
32. Ruiz de Azua I, Scarselli M, Rosemond E, Gautam D, Jou W, Gavrilova O, Ebert PJ, Levitt P, Wess J. RGS4 is a negative regulator of insulin release from pancreatic beta-cells in vitro and in vivo. Proc Natl Acad Sci U S A. 2010;107:7999–8004.
33. Cripps RM, Olson EN. Control of cardiac development by an evolutionarily conserved transcriptional network. Dev Biol. 2002;246:14–28.
34. Kinkel MD, Prince VE. On the diabetic menu: zebrafish as a model for pancreas development and function. BioEssays. 2009;31:139–52.
35. Kimmel RA, Meyer D. Zebrafish pancreas as a model for development and disease. Methods Cell Biol. 2016;134:431–61.
36. Holmstrom SR, Deering T, Swift GH, Poelwijk FJ, Mangelsdorf DJ, Kliewer SA, Macdonald RJ. LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. Genes Dev. 2011;25:1674–9.
37. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjöstedt E, Lundberg E, Szigyarto CA-K, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson A, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13:397–406.
38. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. Nucleic Acids Res. 2012;40:6391–400.
39. Pagliuca FW, Melton DA. How to make a functional β-cell. Development. 2013;140:2472–83.
40. Inagaki N, Kuromi H, Gonoi T, Okamoto Y, Ishida H, Seino Y, Kaneko T, Iwanaga T, Seino S. Expression and role of ionotropic glutamate receptors in pancreatic islet cells. FASEB J. 1995;9:686–91.
41. Villasenor A, Wang ZV, Rivera LB, Ocal O, Asterholm IW, Scherer PE, Brekken RA, Cleaver O, Wilkie TM. Rgs16 and Rgs8 in embryonic endocrine pancreas and mouse models of diabetes. Dis Model Mech. 2010;3:567–80.
42. Zhang W, Morris QD, Chang R, Shai O, Bakowski MA, Mitsakakis N, Mohammad N, Robinson MD, Zirngibl R, Somogyi E, Laurin N, Eftekharpour E, Sat E, Grigull J, Pan Q, Peng W-T, Krogan N, Greenblatt J, Fehlings M, van der Kooy D, Aubin J, Bruneau BG, Rossant J, Blencowe BJ, Frey BJ, Hughes TR. The functional landscape of mouse gene expression. J Biol. 2004;3:21.
43. Hale MA, Swift GH, Hoang CQ, Deering TG, Masui T, Lee Y-K, Xue J, MacDonald RJ. The nuclear hormone receptor family member NR5A2 controls aspects of multipotent progenitor cell formation and acinar differentiation during pancreatic organogenesis. Development. 2014;141:3123–33.
44. Jiang Z, Song J, Qi F, Xiao A, An X, Liu N, Zhu Z, Zhang B, Lin S. Exdpf is a key regulator of exocrine pancreas development controlled by retinoic acid and ptf1a in zebrafish. PLoS Biol. 2008;6:e293.
45. Binot A-C, Manfroid I, Flasse L, Winandy M, Motte P, Martial JA, Peers B, Voz ML. Nkx6.1 and nkx6.2 regulate α- and β-cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. Dev Biol. 2010;340:397–407.
46. Nica AC, Ongen H, Irminger J-CC, Bosco D, Berney T, Antonarakis SE, Halban PA, Dermitzakis ET. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. Genome Res. 2013;23:1554–62.
47. Devos N, Deflorian G, Biemar F, Bortolussi M, Martial JA, Peers B, Argenton F. Differential expression of two somatostatin genes during zebrafish embryonic development. Mech Dev. 2002;115:133–7.
48. Blodgett DM, Nowosielska A, Afik S, Pechhold S, Cura AJ, Kennedy NJ, Kim S, Kucukural A, Davis RJ, Kent SC, Greiner DL, Garber MG, Harlan DM, Dilorio P. Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. Diabetes. 2015;64:3172–81.
49. Tian G, Sandler S, Gylfe E, Tengholm A. Glucose- and hormone-induced cAMP oscillations in α- and β-cells within intact pancreatic islets. Diabetes. 2011;60:1535–43.
50. Davidson AJ, Ernst P, Wang Y, Dekens MPS, Kingsley PD, Palis J, Korsmeyer SJ, Daley GQ, Zon LI. cdx4 mutants fail to specify blood progenitors and can be rescued by multiple hox genes. Nature. 2003;425:300–6.
51. Kinkel MD, Eames SC, Alonzo MR, Prince VE. Cdx4 is required in the endoderm to localize the pancreas and limit beta-cell number. Development. 2008;135:919–29.
52. Bertrand G, Gross R, Puech R, Loubatières-Mariani MM, Bockaert J. Evidence for a glutamate receptor of the AMPA subtype which mediates insulin release from rat perfused pancreas. Br J Pharmacol. 1992;106:354–9.
53. Lee EK, Kim W, Tominaga K, Martindale JL, Yang X, Subaran SS, Carlson OD, Mercken EM, Kulkarni RN, Akamatsu W, Okano H, Perrone-Bizzozero NI, de Cabo R, Egan JM, Gorospe M. RNA-binding protein HuD controls insulin translation. Mol Cell. 2012;45:826–35.
54. Mergler S, Singh V, Grötzinger C, Kaczmarek P, Wiedenmann B, Strowski MZ. Characterization of voltage operated R-type Ca2+ channels in modulating

somatostatin receptor subtype 2- and 3-dependent inhibition of insulin secretion from INS-1 cells. Cell Signal. 2008;20:2286–95.
55. Squires PE, Jones PM, Younis MYG, Hills CE. Chapter Ten – The Calcium-Sensing Receptor and β-Cell Function. In: Litwack G, editor. Vitamins & Hormones, The Pancreatic Beta Cell, vol. 95. Cambridge: Academic Press; 2014. p. 249–67.
56. Cortijo C, Gouzi M, Tissir F, Grapin-Botton A. Planar cell polarity controls pancreatic beta cell differentiation and glucose homeostasis. Cell Rep. 2012;2:1593–606.
57. Flanagan SE, De Franco E, Lango Allen H, Zerah M, Abdul-Rasoul MM, Edge JA, Stewart H, Alamiri E, Hussain K, Wallis S, de Vries L, Rubio-Cabezas O, Houghton JAL, Edghill EL, Patch A-M, Ellard S, Hattersley AT. Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. Cell Metab. 2014;19:146–54.
58. Jennings RE, Berry AA, Strutt JP, Gerrard DT, Hanley NA. Human pancreas development. Development. 2015;142(18):3126–37.
59. Kilic G, Alvarez-Mercado AI, Zarrouki B, Opland D, Liew CW, Alonso LC, Myers MG, Jonas J-C, Poitout V, Kulkarni RN, Mauvais-Jarvis F. The islet estrogen receptor-α is induced by hyperglycemia and protects against oxidative stress-induced insulin-deficient diabetes. PLoS One. 2014;9(2):e87941.
60. Yuchi Y, Cai Y, Legein B, De Groef S, Leuckx G, Coppens V, Van Overmeire E, Staels W, De Leu N, Martens G. Estrogen receptor α regulates beta cell formation during pancreas development and following injury. Diabetes. 2015;64(9):3218–28. doi:10.2337/db14-1798.
61. Ye L, Robertson MA, Hesselson D, Stainier DYR, Anderson RM. glucagon is essential for alpha cell transdifferentiation and beta cell neogenesis. Development. 2015;142:1407–17.
62. Chera S, Baronnier D, Ghila L, Cigliola V, Jensen JN, Gu G, Furuyama K, Thorel F, Gribble FM, Reimann F, Herrera PL. Diabetes recovery by age-dependent conversion of pancreatic δ-cells into insulin producers. Nature. 2014;514:503–7.
63. Courtney M, Gjernes E, Druelle N, Ravaud C, Vieira A, Ben-Othman N, Pfeifer A, Avolio F, Leuckx G, Lacas-Gervais S, Burel-Vandenbos F, Ambrosetti D, Hecksher-Sorensen J, Ravassard P, Heimberg H, Mansouri A, Collombat P. The inactivation of Arx in pancreatic α-cells triggers their neogenesis and conversion into functional β-like cells. PLoS Genet. 2013;9:e1003934.
64. Gao T, McKenna B, Li C, Reichert M, Nguyen J, Singh T, Yang C, Pannikar A, Doliba N, Zhang T, Stoffers DA, Edlund H, Matschinsky F, Stein R, Stanger BZ. Pdx1 maintains β-cell identity and function by repressing an α-cell program. Cell Metab. 2014;19:259–71.
65. Thorel F, Nepote V, Avril I, Kohno K, Desgraz R, Chera S, Herrera PL. Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature. 2010;464:1149–54.
66. Spijker HS, Ravelli RBG, Mommaas-Kienhuis AM, van Apeldoorn AA, Engelse MA, Zaldumbide A, Bonner-Weir S, Rabelink TJ, Hoeben RC, Clevers H, Mummery CL, Carlotti F, de Koning EJP. Conversion of mature human β-cells into glucagon-producing α-cells. Diabetes. 2013;62:2471–80.
67. Wendik B, Maier E, Meyer D. Zebrafish mnx genes in endocrine and exocrine pancreas formation. Dev Biol. 2004;268:372–83.
68. Pan FC, Brissova M, Powers AC, Pfaff S, Wright CVE. Inactivating the permanent neonatal diabetes gene Mnx1 switches insulin-producing β-cells to a δ-like fate and reveals a facultative proliferative capacity in aged β-cells. Development. 2015;142:3637–48.
69. Nelson SB, Schaffer AE, Sander M. The transcription factors Nkx6.1 and Nkx6.2 possess equivalent activities in promoting beta-cell fate specification in Pdx1+ pancreatic progenitor cells. Development. 2007;134:2491–500.
70. Flasse LC, Pirson JL, Stern DG, Von Berg V, Manfroid I, Peers B, Voz ML. Ascl1b and Neurod1, instead of Neurog3, control pancreatic endocrine cell fate in zebrafish. BMC Biol. 2013;11:78.
71. Flasse LC, Stern DG, Pirson JL, Manfroid I, Peers B, Voz ML. The bHLH transcription factor Ascl1a is essential for the specification of the intestinal secretory cells and mediates Notch signaling in the zebrafish intestine. Dev Biol. 2013;376:187–97.
72. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, Stainier DYR. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015;524:230–3.
73. McIntyre DC, Rakshit S, Yallowitz AR, Loken L, Jeannotte L, Capecchi MR, Wellik DM. Hox patterning of the vertebrate rib cage. Development. 2007;134(16):2981–9.
74. Hummler E, Cole TJ, Blendy JA, Ganss R, Aguzzi A, Schmid W, Beermann F, Schütz G. Targeted mutation of the CREB gene: compensation within the CREB/ATF family of transcription factors. Proc Natl Acad Sci U S A. 1994;91:5647–51.
75. Wang S, Zhang J, Zhao A, Hipkens S, Magnuson MA, Gu G. Loss of Myt1 function partially compromises endocrine islet cell differentiation and pancreatic physiological function in the mouse. Mech Dev. 2007;124:898–910.
76. Wang S, Hecksher-Sorensen J, Xu Y, Zhao A, Dor Y, Rosenberg L, Serup P, Gu G. Myt1 and Ngn3 form a feed-forward expression loop to promote endocrine islet cell differentiation. Dev Biol. 2008;317:531–40.
77. Tennant BR, Islam R, Kramer MM, Merkulova Y, Kiang RL, Whiting CJ, Hoffman BG. The transcription factor Myt3 acts as a pro-survival factor in β-cells. PLoS One. 2012;7:e51501.
78. Brereton H, Carvell MJ, Persaud SJ, Jones PM. Islet alpha-cells do not influence insulin secretion from beta-cells through cell-cell contact. Endocrine. 2007;31:61–5.
79. Ohta Y, Kosaka Y, Kishimoto N, Wang J, Smith SB, Honig G, Kim H, Gasa RM, Neubauer N, Liou A, Tecott LH, Deneris ES, German MS. Convergence of the insulin and serotonin programs in the pancreatic β-cell. Diabetes. 2011;60:3208–16.
80. Kwan KM, Fujimoto E, Grabher C, Mangum BD, Hardy ME, Campbell DS, Parant JM, Yost HJ, Kanki JP, Chien CB. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Dev Dyn. 2007;236:3088–99.
81. Zecchin E, Filippi A, Biemar F, Tiso N, Pauls S, Ellertsdottir E, Gnügge L, Bortolussi M, Driever W, Argenton F. Distinct delta and jagged genes control sequential segregation of pancreatic cell types from precursor pools in zebrafish. Dev Biol. 2007;301:192–204.
82. Ramsköld D, Luo S, Wang Y-C, Li R, Deng Q, Faridani OR, Daniels GA, Khrebtukova I, Loring JF, Laurent LC, Schroth GP, Sandberg R. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol. 2012;30:777–82.
83. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–78.
84. Moss SP, Joyce DA, Humphries S, Tindall KJ, Lunt DH. Comparative analysis of teleost genome sequences reveals an ancient intron size expansion in the zebrafish lineage. Genome Biol Evol. 2011;3:1187–96.
85. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9. doi:10.1093/bioinformatics/btu638.
86. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.
87. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
88. Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E. EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Res. 2009;19:327–35.
89. Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, Almeida-King J, Staines D, Derwent P, Kerhornou A, Kersey P, Flicek P. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford). 2011;2011:bar030. doi:10.1093/database/bar030.
90. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2008;4:44–57.
91. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc. 2008;3:59–69.
92. Lillesaar C, Tannhäuser B, Stigloher C, Kremmer E, Bally-Cuif L. The serotonergic phenotype is acquired by converging genetic mechanisms within the zebrafish central nervous system. Dev Dyn. 2007;236:1072–84.
93. Mavropoulos A, Devos N, Biemar F, Zecchin E, Argenton F, Edlund H, Motte P, Martial JA, Peers B. sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. Dev Biol. 2005;285:211–23.
94. Jao L-EE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A. 2013;110:13904–9.
95. Varshney GK, Pei W, LaFave MC, Idol J, Xu L, Gallardo V, Carrington B, Bishop K, Jones M, Li M, Harper U, Huang SC, Prakash A, Chen W, Sood R, Ledin J, Burgess SM. High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. Genome Res. 2015;25:1030–42.

1.1 Overview of the zebrafish RNA-Seq data

We actively continued to use our zebrafish bulk RNA-Seq data from major pancreatic cell types to perform further analyses. This section presents a brief summary and an update of the previous results, also covering different aspects not highlighted in the publication.

We performed an update of our analyses by using a new pipeline and a more recent version of the reference genome and its annotations (see methods 4.3). As presented in the paper, each library was firstly evaluated by looking at their respective expression of key genes specific of each pancreatic population. Hormone genes were evaluated, and results showed they were the highest represented transcripts in each corresponding endocrine cell types. *Gcga* and *gcgb* transcripts represent 16.77% (± 0.66) of all reads in alpha samples (Table 1). *Ins* transcripts cover 16.32% (± 5.19) of beta samples while *sst2* and *sst1.2* share 33.37% (± 8.86) of all reads in delta samples. Samples for ductal cells display a different distribution as they don't strongly express a specific gene. Two of the most expressed genes, *tmsb4x* and *actb2*, cover together only 2.53% (± 0.32) of all reads. They are involved in cytoskeleton organization. Finally, acinar samples have 6 genes coding for digestive enzymes whose combined transcripts cover 44.46% (± 1.26) of reads.

These data confirm the good purity of each sample as specific markers of other cell types were not significantly detected (previously described in the paper). We also compared all RNA-seq data sets by performing Principal Component Analysis (PCA) and clustering analysis. This first highlighted a strong separation between endocrine, acinar and ductal samples. This observation confirms that endocrine cells present a more similar transcriptome between them compared to ductal or acinar cells. These two are both exocrine-associated cell types but they present significative transcriptomic differences underlying their distinct functions. By focusing the analysis on endocrine cells, we highlighted a clear distinction between alpha, beta and delta samples (Fig 25).

Name	Alpha	Beta	Delta	Acinar	Ductal

Table 1: Most expressed genes by cell type. Expressions are displayed in percentage of total reads detected by sample.

Name	Alpha	Beta	Delta	Acinar	Ductal
gcga	12,49	0,17	0,32	0,00	0,00
gcgb	4,29	0,02	0,14	0,00	0,00
ins	0,11	16,32	0,07	0,00	0,05
sst2	0,14	0,04	19,53	0,00	0,00
sst1.2	0,10	0,04	13,85	0,00	0,00
prss1	0,08	0,00	0,07	9,24	0,08
cpa5	0,07	0,00	0,05	8,12	0,04
prss59.1	0,07	0,00	0,06	8,39	0,12
prss59.2	0,06	0,00	0,05	7,93	0,11
cpb1	0,06	0,00	0,04	5,63	0,03
ela2l	0,06	0,00	0,06	5,20	0,09
tmsb4x	0,02	0,05	0,03	0,01	1,24
actb2	0,14	0,21	0,15	0,18	1,28

Using RNA-Seq, we can explore the transcriptome by looking at different aspects of the data. Our first approach was to consider the set of genes which present an enriched expression in each cell type as they highlight the transcriptional specificity. Based on our data, these lists of enriched genes for each population has been updated and we set up lists of **47, 57, 172, 842, 634** enriched genes for alpha, beta, delta, ductal and acinar cells, respectively (see methods 4.3). We also highlighted 482 and 133 genes with 10 times enrichment in endocrine (alpha, beta, delta) and exocrine (ductal, acinar) tissue, respectively (Fig 26 + Supplemental table 1).

The most common approach with RNA-Seq is to define a set of genes with enriched expression in one condition compared to another. This method is dependent of the set

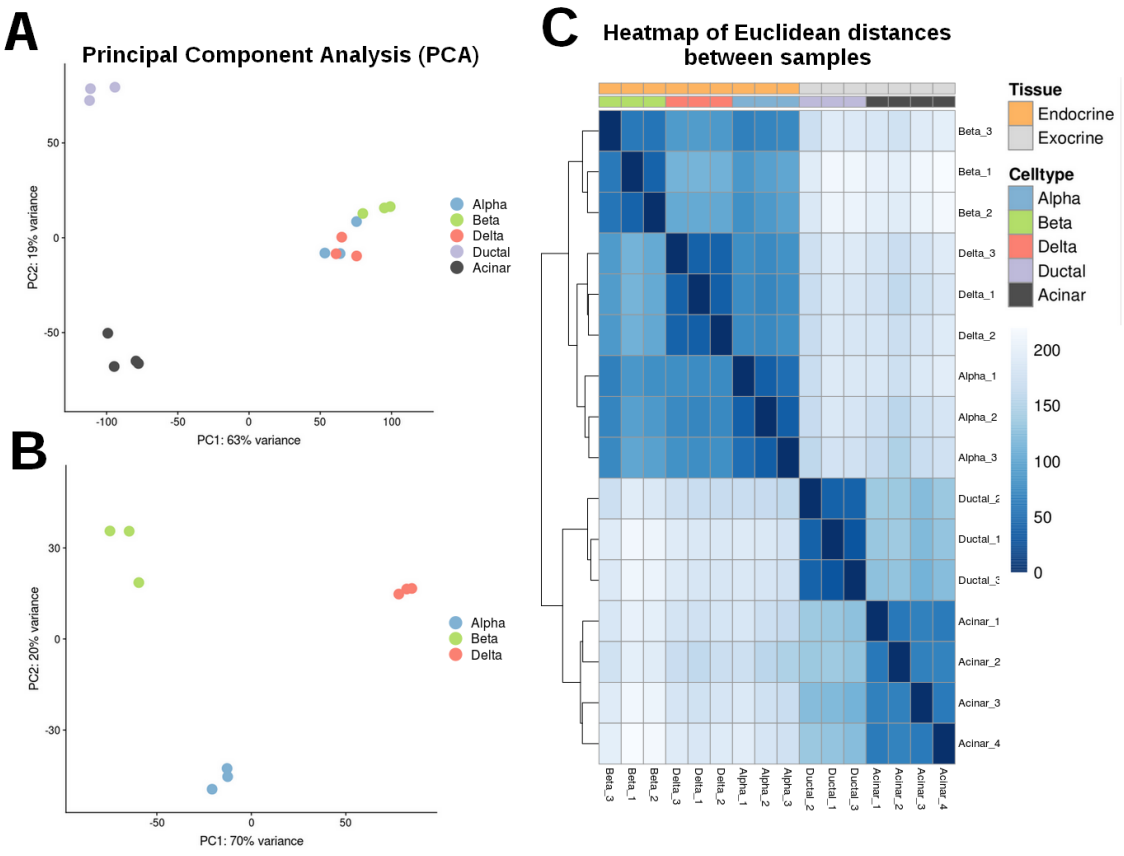


Figure 25: **Overview of the analysis of data**. (A) PCA of the five pancreatic cell types which highlights separation between endocrine, ductal and acinar cells. (B) PCA focusing on the 3 endocrine cell subtypes which shows distinct clusters. (C) Hierarchical clustering of all the samples.

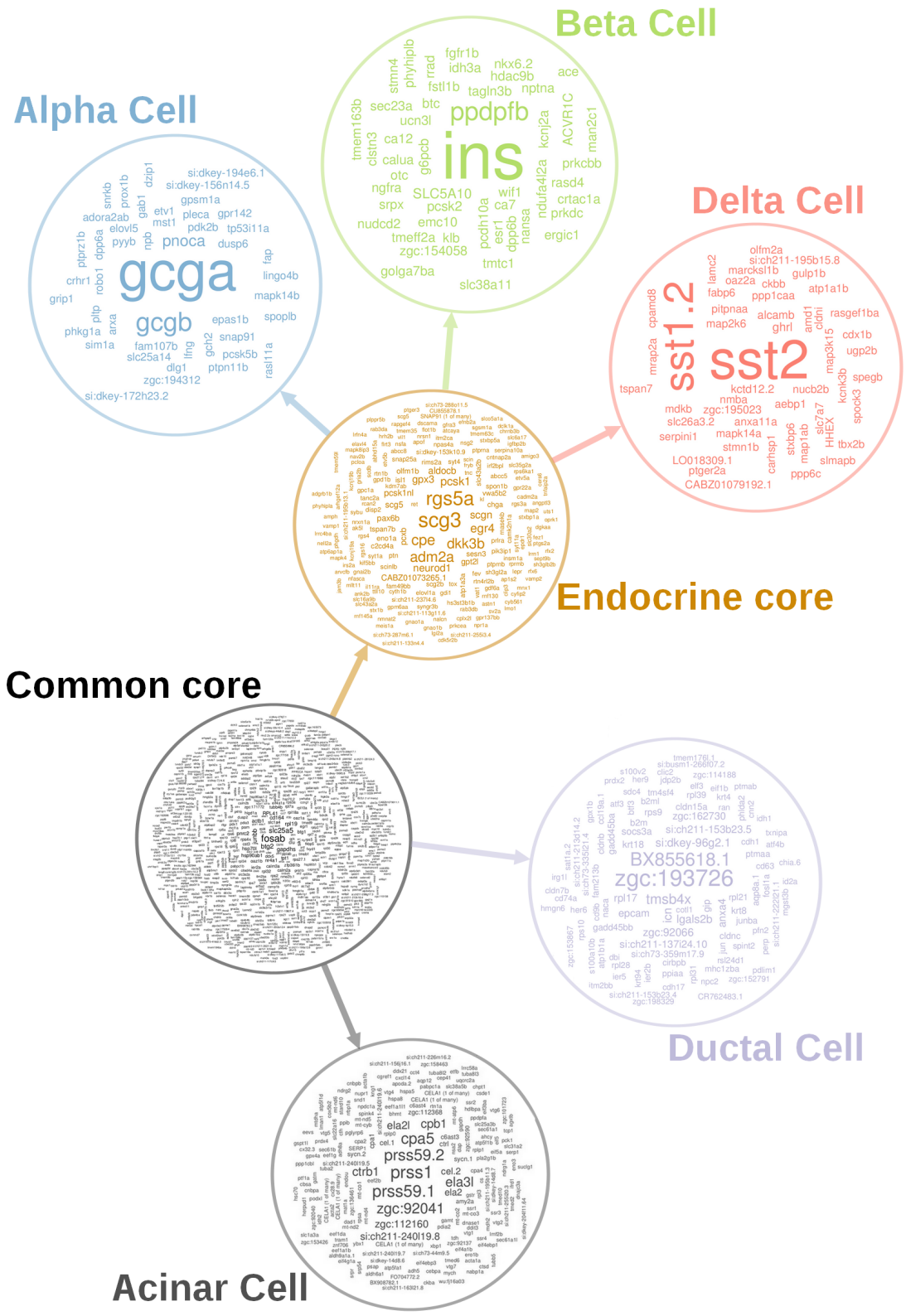
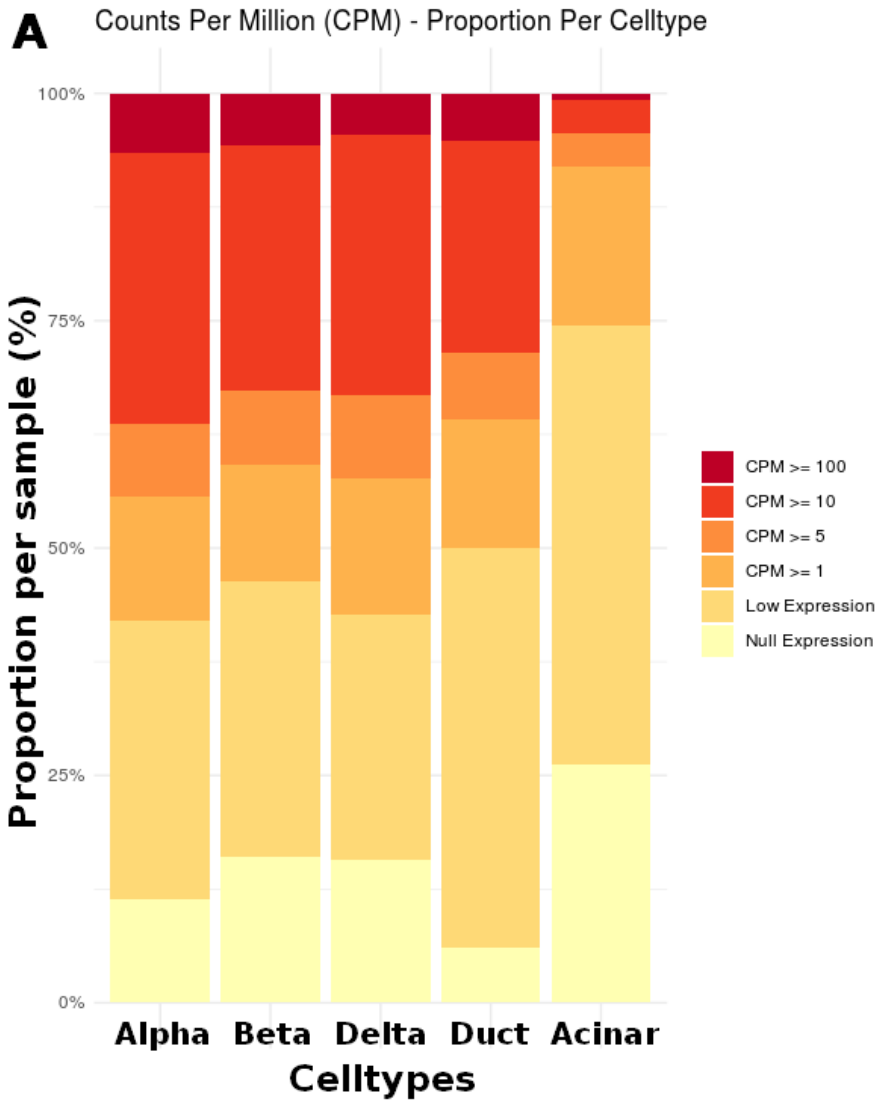


Figure 26: **Expression profile of genes in zebrafish bulk RNA-seq data**. Each group contains genes with a significative enriched expression for this group. The fontsize of the gene names represents their relative expression levels.

of different conditions that we have at our disposal. When comparing cell types, these subsets of genes represent potential marker genes of each cell type. The accuracy of these lists is directly correlated with the amount of cell types which are considered.
The transcriptomic signature of a cell type is not only restricted to the set of enriched markers. It represents the global set of RNAs found in a cell and by extension regarding at mRNA, the global set of genes expressed. Using a reference set of 25106 protein coding genes (Ensembl release 92), we evaluated the transcriptomic signature of the five pancreatic cell types. We detected for endocrine cells 14565 (58.01%) genes expressed in alpha cells (Counts Per Million (CPM) $\geq$ 1), 13450 (53.57%) genes in beta cells and 14398 (57.34%) genes in delta cells. For exocrine tissue, we detected 12534 (49.92%) and 6421 (25.57%) genes expressed in ductal and acinar cells, respectively. Endocrine and ductal cells have similar distribution of expression levels: 5.52% (± 0.85) of genes with CPM $\geq$ 100, 27.12% (± 2.80) of CPM $\geq$10, 8.21% (± 0.79) of CPM $\geq$ 5 and 13.85 % (± 0.89) of genes with CPM $\geq$ 1 (Fig 27). The acinar samples present a lower complexity meaning that many sequenced reads correspond to the same molecules implying a level of redundancy in the data [295]. This affects the distribution of expression levels which is therefore different from other cell types: 0.72% of genes with CPM $\geq$ 100, 3.60% of CPM $\geq$ 10, 3.80% of CPM $\geq$ 5 and 17.45% of CPM $\geq$ 1. This means that most of the reads cover very few genes and implies that in some extend these libraries are less balanced than the others.

As our previous results showed that each pancreatic endocrine cell type does not express many specific genes when compared one to another, it implies that they should have high similarities between them. We decided to quantify this by comparing their global transcriptomes. We considered total expressed genes (CPM > 1) and we observed that around 85% (87.18 ± 3.79) of genes expressed by one endocrine cell type is also expressed in the two others (Figure 28). This value rises to around 95% (95.23 ± 2.04) for genes expressed in at least two endocrine cell types. We then estimated the overlap between genes expressed in ductal cells and genes expressed in each endocrine cell type taken separately. In comparison, we observed an average overlap around 75% (76.56 ± 1.69) between any endocrine cell and ductal cell transcriptomes.

In conclusion, we described each pancreatic cell type by defining sets of representative



B

	CPM_100	CPM_10	CPM_5	CPM_1	CPM_Low	CPM_Null
Alpha	6.56	29.76	8.07	13.63	30.53	11.46
Beta	5.79	26.88	8.11	12.79	30.39	16.04
Delta	4.57	28.55	9.29	14.93	26.92	15.73
Ductal	5.18	23.31	7.37	14.06	44.03	6.04
Acinar	0.72	3.60	3.80	17.45	48.22	26.20

Figure 27: **General comparison of transcriptomes**. (A) Distribution of genes according to their CPM levels. (B) Values associated. It shows that acinar samples present a different profile, probably due to a lower complexity of the library.

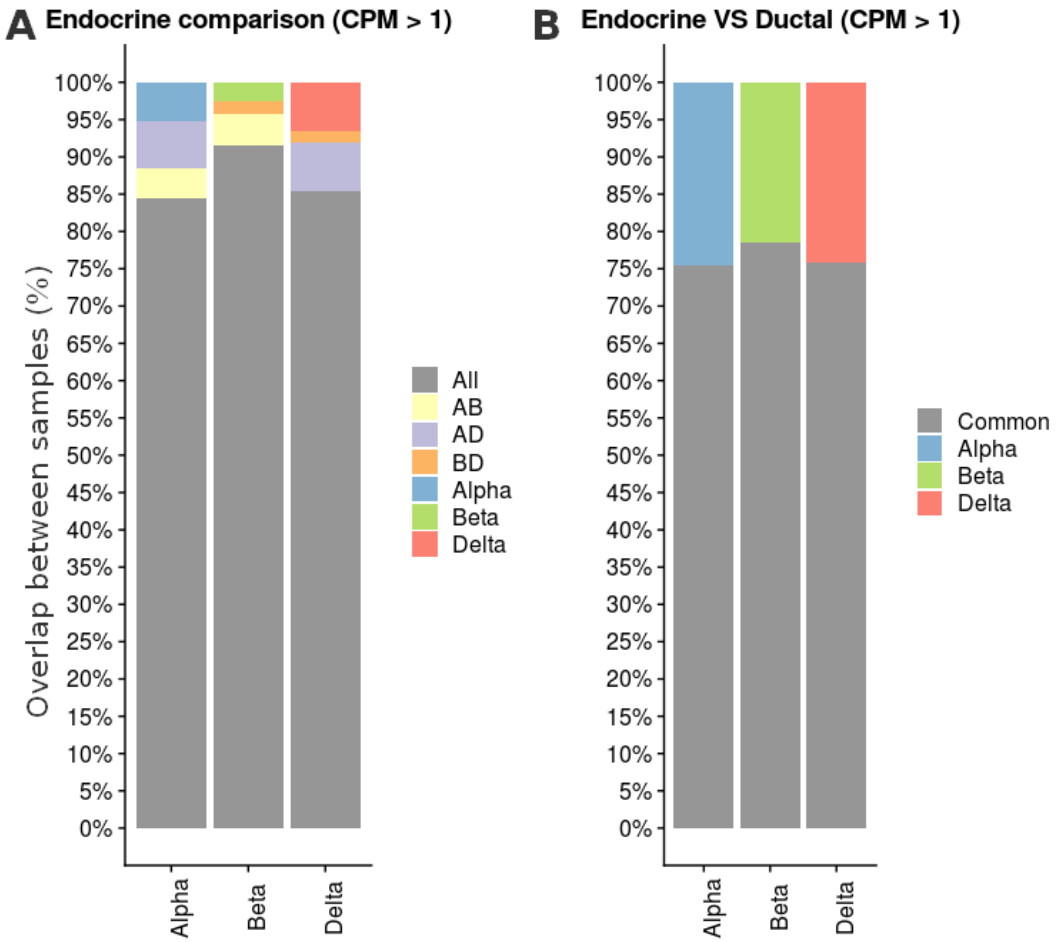


Figure 28: **Comparison of endocrine transcriptomes**. (A) The transcriptomes of endocrine cells present a strong overlap between them. Around 95% of all genes expressed (CPM > 1) in an endocrine cell subtype is expressed in at least 1 of the others. AB = expressed in alpha and beta; AD = expressed in alpha and delta; BD = expressed in beta and delta. (B) When comparing the transcriptome of each endocrine cell subtype separately with the transcriptome of ductal cell, the overlap decreases to ~75%.

genes for each but also by the complete set of expressed genes. We highlighted a close similarity between the three endocrine cell signatures.

1.2 Zebrafish scRNA-Seq from adult pancreatic tissues

In order to update and complete our atlas of all pancreatic cell transcriptomes, we decided to use available Single-Cell RNA-Seq data produced in adult pancreatic tissues from zebrafish by the group of Ninov N. [147]. We retrieved the adult pancreatic datasets from their study and we reanalyzed them to determine if they fit with our previous bulk RNA-Seq work. After multiple filtering steps (see Methods 4.5.1), we kept an amount of 6515 single cells to use with average 1222 (± 496) expressed genes detected per cell. Cell clustering highlighted 7 major groups of pancreatic cells based on transcriptomic signal. We evaluate expression pattern of key genes across the different clusters in order to determine their pancreatic cell identities (Figure 29-A). As with our bulk RNA-Seq we only isolated 5 pancreatic cell types using specific transgenic lines (alpha, beta, delta, ductal and acinar cells), scRNA-Seq data present two more clusters. One of these is characterized by the specific expression of *sst1.1* gene, that we refer as delta 1.1 cells. The second cluster correspond to ghrelin-expressing epsilon cells. Using combinations of specific markers for each group, we created "cell-score" to define and annotate accurately each pancreatic cell population. After this process, we obtained the following distribution of cells: 618 alpha cells, 1383 beta cells, 2255 delta (sst2) cells, 746 delta (sst1.1) cells, 91 epsilon cells, 875 ductal cells and 547 acinar cells (Figure 29-B).

To evaluate this annotation, we decided to screen expression of pancreatic cell type markers that we determined in our bulk RNA-Seq in order to estimate in which proportion they can also be validated with the single-cell (SC) data. We observed that a large majority of these markers were assigned to the same cell type (Figure 30-A). But as two additional pancreatic cell transcriptomes are detected with the SC data, some genes need to be reassigned to appropriate cell type. This is the case for genes such as *pcsk2* and *spon1b* that we previously assigned to beta cells but seem to be preferentially expressed in delta 1.1 cells. Similarly, *pdx1* which is vastly described as beta marker but seems to be highly expressed in delta 1.1 cells in zebrafish.

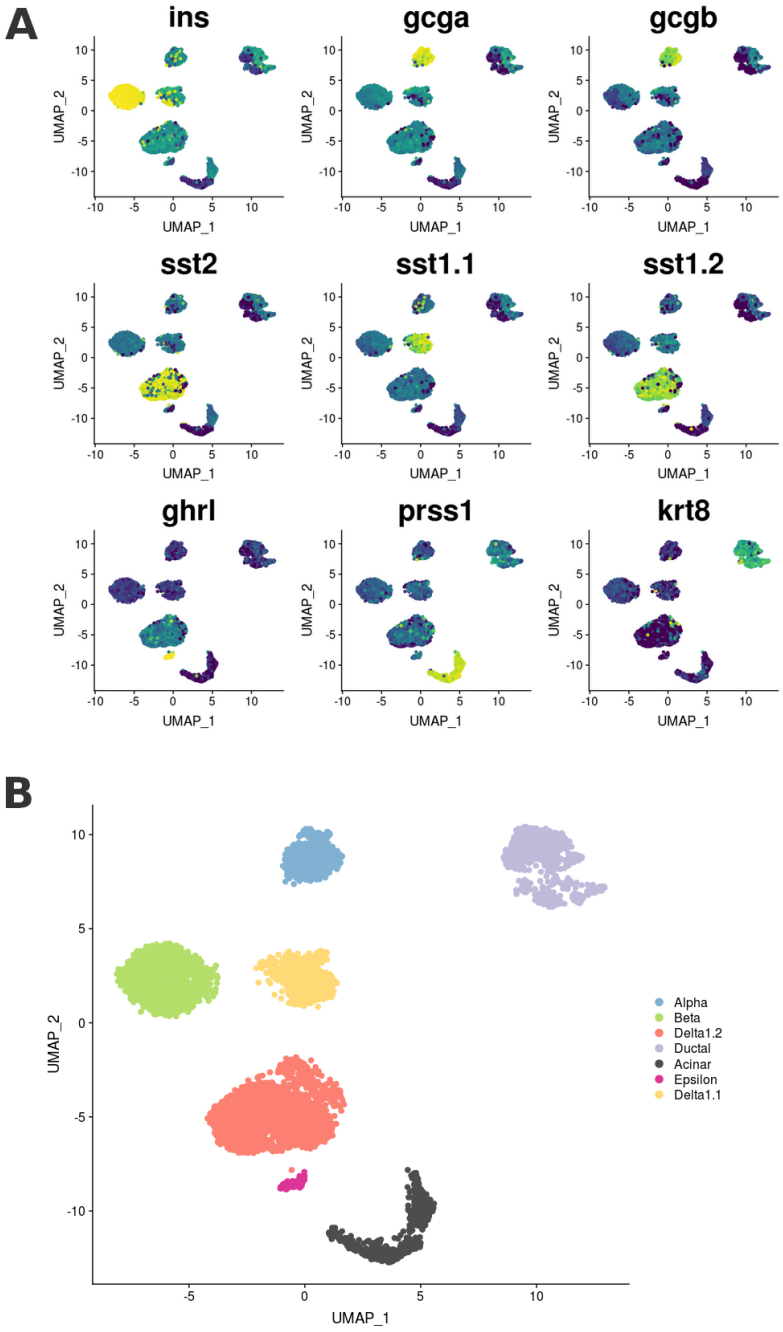


Figure 29: **Analysis of Zebrafish scRNA-Seq data**. (A) Expression of key genes of each cell type on the UMAP embedding from SC data. (B) Annotation of the SC clusters with the corresponding cell type.

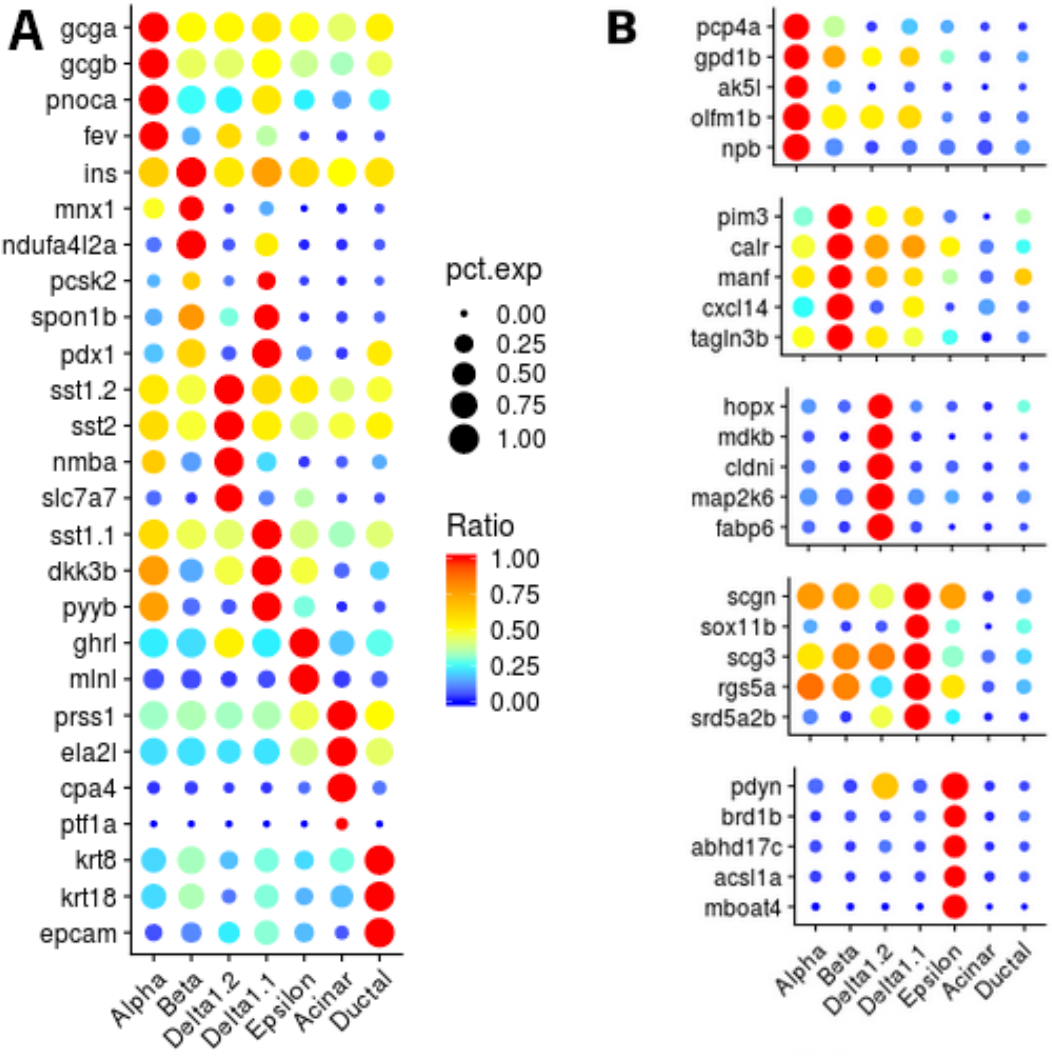


Figure 30: **Validation of marker genes**. (A) DotPlot of previously described marker genes from our bulk RNA-seq data. Some genes presented enrichment for the two additionnal signature. (B) DotPlot of endocrine cells marker genes identified from the scRNA-seq data. Size of the dot represents the proportion of cells in a cluster expressing the gene. Color of the dot represents the relative expression level of the gene in a celltype compared to the celltype with the highest expression.

These observations indicate that the SC data may complement our bulk data in defining cell type specific gene signatures. We started by determining all genes expressed in each cluster before performing the differential expression analysis. In single-cell experiments, only a small part of the global transcriptome is detected for a cell. After the identification of the different clusters, each cell contributes to reconstruct the global transcriptome of the cell type to which it belongs. The amount of cells in a cluster impacts the covering of the associated transcriptome. We observed an average of 13681 (± 1803) expressed genes detected by cluster which is consistent with the average of 12273 (±3371) genes that we detected with bulk data. We determined new lists of markers for each pancreatic cell type according to SC particularity. Indeed, the differential expression of a gene can be related to two parameters: the percentage of cells expressing the gene in each cluster, and the expression level of the gene in both clusters (Figure 31).

After differential expression analysis we identified 22, 30, 116, 168, 20, 180 and 61 enriched genes for alpha, beta, delta 1.2, delta 1.1, epsilon, ductal and acinar cells, respectively (with the threshold of being expressed in at least 50% of cells of a cluster in order to be considered as marker of this group). We determined also a set of 227 endocrine enriched genes (Figure 32 + Supplemental table 2). Similarly to our bulk data, for each endocrine cell type, the associated hormone is found in the marker list. Concerning acinar cells, we found multiple genes coding for digestive enzymes. Results also validate some markers we previously identified such as *pnoca* and *fev* for alpha cells, *ndufa4l2a* and *tagln3b* for beta cells or *nmba* and *slc7a7* for delta 1.2 cells. We also defined markers for epsilon cells with genes such as *mboat4*, *acsl1a*, *abhd17c* and *brd1b* which seem highly specific. The first three are known to play a role in fatty acid metabolism with *acsl1a* coding for an enzyme involved in the conversion of long-chain fatty acids into acyl-CoA. Acyl-CoA is required for ghrelin acylation with *mboat4* encoding the acyltransferase responsible for this process and also responsible for the octanoylation of ghrelin at 'Ser-3' [296, 297]. *Mlnl* expression is strongly enriched in epsilon cells and this hormone is known to be strongly expressed in enteroendocrine cells in the gut. Finally, as we obtained the transcriptome of delta 1.1 cells, we highlighted unknown markers for this population mainly defined by the specific expression of *sst1.1* gene. It is also interesting to observe that some genes previously associated with other cell types seem now more enriched

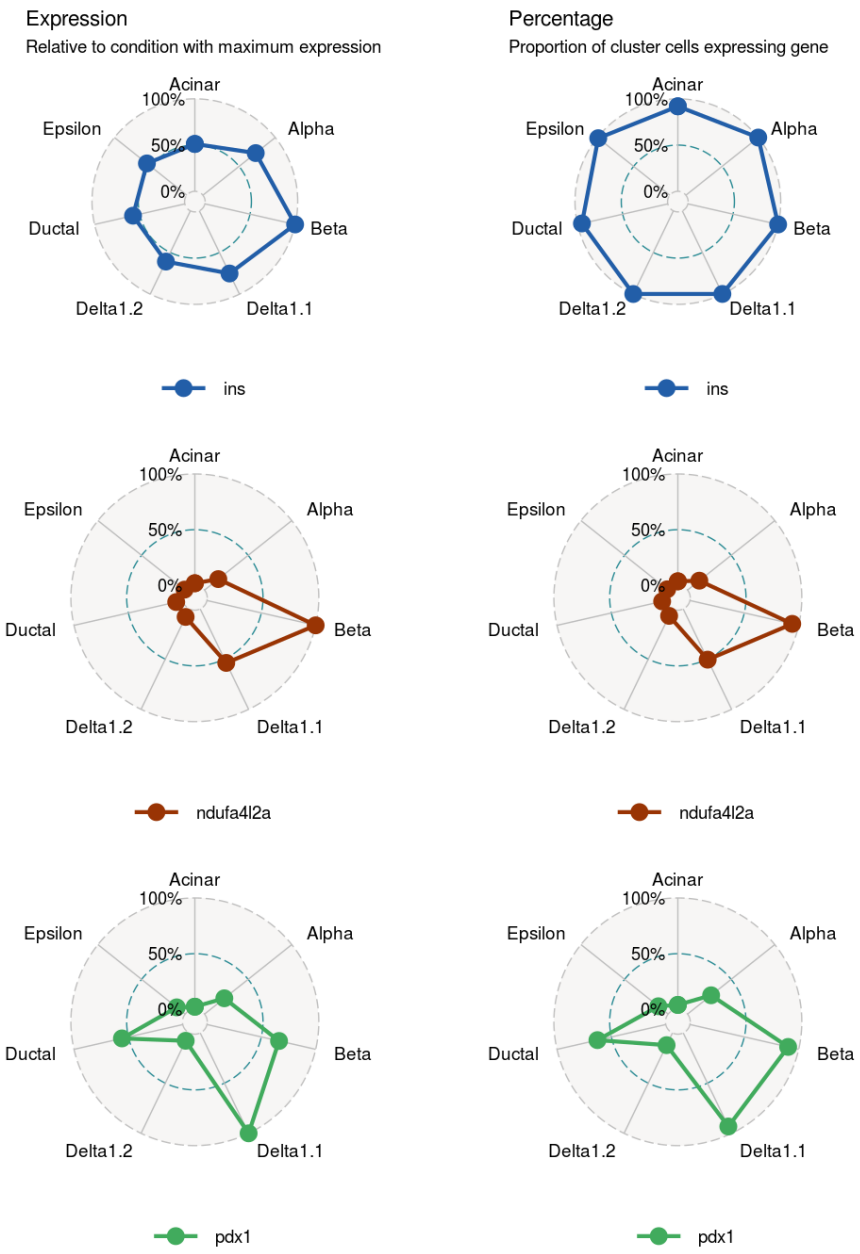


Figure 31: **Characteristics of scRNA-seq data**. Two parameters are important with SC data: the expression level and the percentage of cells expressing the gene in a cluster. Genes highly expressed such as *ins* gene, are detected nearly in all cells by SC but the expression level allows to define the correct profile of expression. *ndufa4l2a* profile is easily identifiable as both parameters indicate the same cell type. *pdx1* profile is interesting as two cell types present the same detection percentage with only the expression level separate them.

in this group such as *dkk3b* or *pyyb* (Figure 30-A). As mentioned previously, the most surprising result is *pdx1* expression which is enriched in delta 1.1 cells compared to beta cells (~1.5x). *Pdx1* has been widely described as a beta identity gene (notably in mammals) with a key role in the regulation of insulin gene transcription. Interestingly, PDX1 was initially identified as a transcription factor involved in the activation of somatostatin gene [298]. This raises questions about the role of this subpopulation in zebrafish pancreas as a similar subpopulation of "delta" cells is not described in mouse or human pancreas.

1.3 Atlas of endocrine pancreatic cell transcriptomes in Zebrafish

Based on the results obtained in our bulk RNA-seq and in available scRNA-seq, we defined set of key marker genes for each endocrine cell type in zebrafish. Hormone coding genes are obviously strong markers to identify each endocrine cell type. For beta cells, genes such as *ndufa4l2a*, *cxcl14*, *mnx1*, *ppdpfb*, *nkx6.2* and *tagln3b* showed enrichment in both experiments and their concomitant expression can be used to recognize beta-cell identity in pancreas (Figure 33). We decided not to include *pdx1* as its expression is significantly higher in delta 1.1 cells. Beside both glucagon genes, we can use *pnoca*, *npb*, *etv1*, *scinlb*, *fev* and *olfm1b* expression as an alpha identity signature in zebrafish. *Arxa* is generally associated with alpha cell type but our results showed a strong expresion in epsilon cells as well and we decided not to keep it for the alpha cell identity. For the delta 1.2 cell identity, somatostatin coding genes *sst2* and *sst1.2* are the strongest markers but we also added *map2k6*, *nmba*, *slc7a7*, *map3k15*, *cdx4* and *lamc2* genes to define the signature. Signature for epsilon identity could only be defined based on scRNA-seq results as we didn't produce RNA-seq for this cell type in the laboratory. The highest markers of epsilon cells are the hormone coding genes *ghrl* and *mlnl* to which we also added *mboat4*, *acsl1a*, *abhd17c*, *brd1b* and *plcxd3*. Finally, the most difficult signature to define was for delta 1.1 cells. These cells express many genes also with enriched expression in another endocrine cell subtypes such as *pdx1*, *dkk3b* or *pyyb*. Nevertheless, 4 genes display specific expression in delta 1.1 cells: *sst1.1*, *sox11b*, *nsg2* and *srd5a2b*.

As we did with bulk RNA-seq, we compared the whole transcriptomes of each endocrine

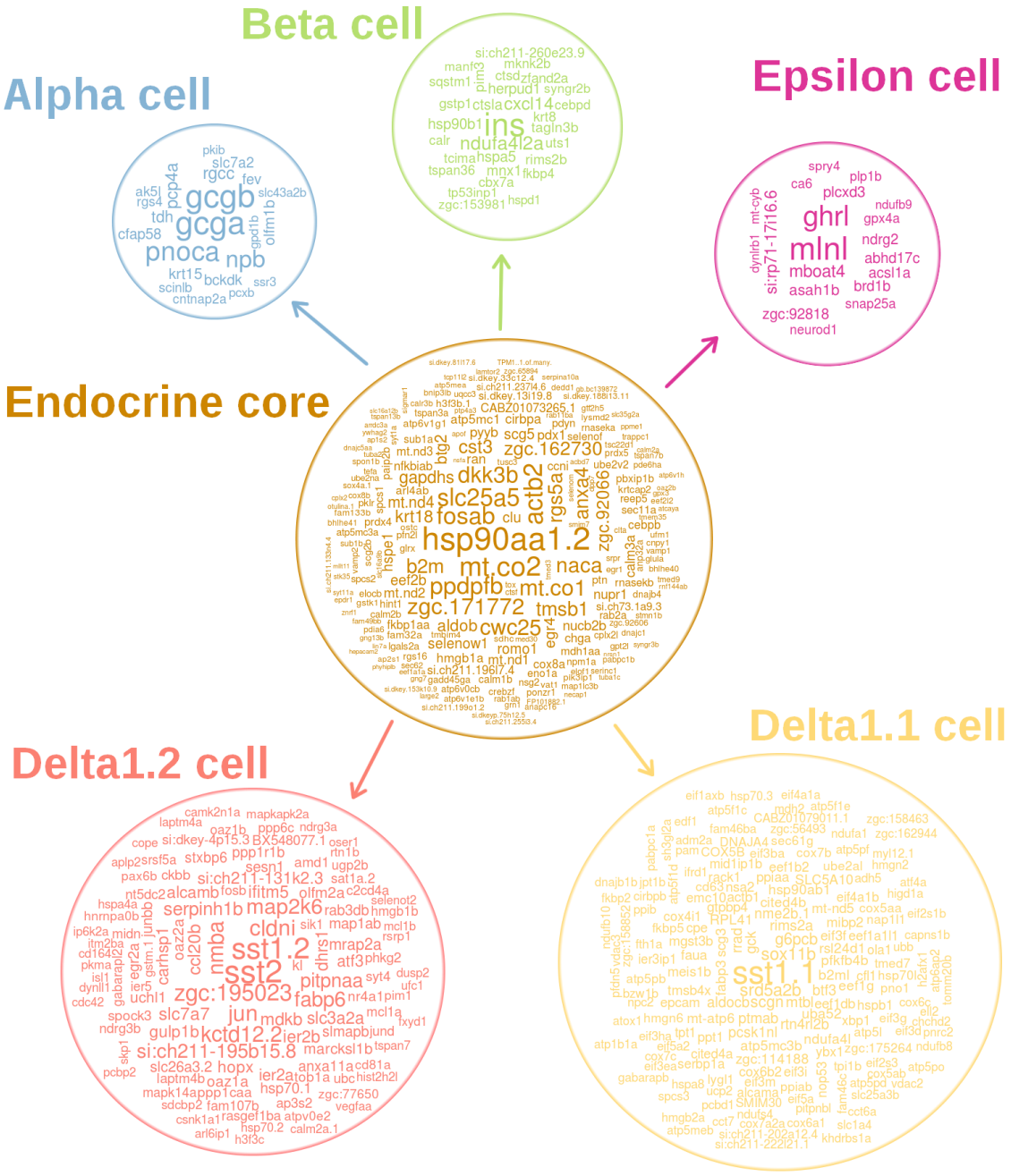


Figure 32: **Expression profile of genes in zebrafish SC data**. Each group contains genes with a significative enriched expression for this group. The size of the gene names represents their relative expression levels.

cell type and we observed that more than 85 % of expressed genes by one endocrine cell type are also expressed by all other endocrine cell types (excluding epsilon signal as low number of cells only partially cover transcriptome). On the opposite, less than 2% of expressed genes in an endocrine cell type are only expressed in this cell type. These results highlight the close similarity of transcriptomes between endocrine cells.

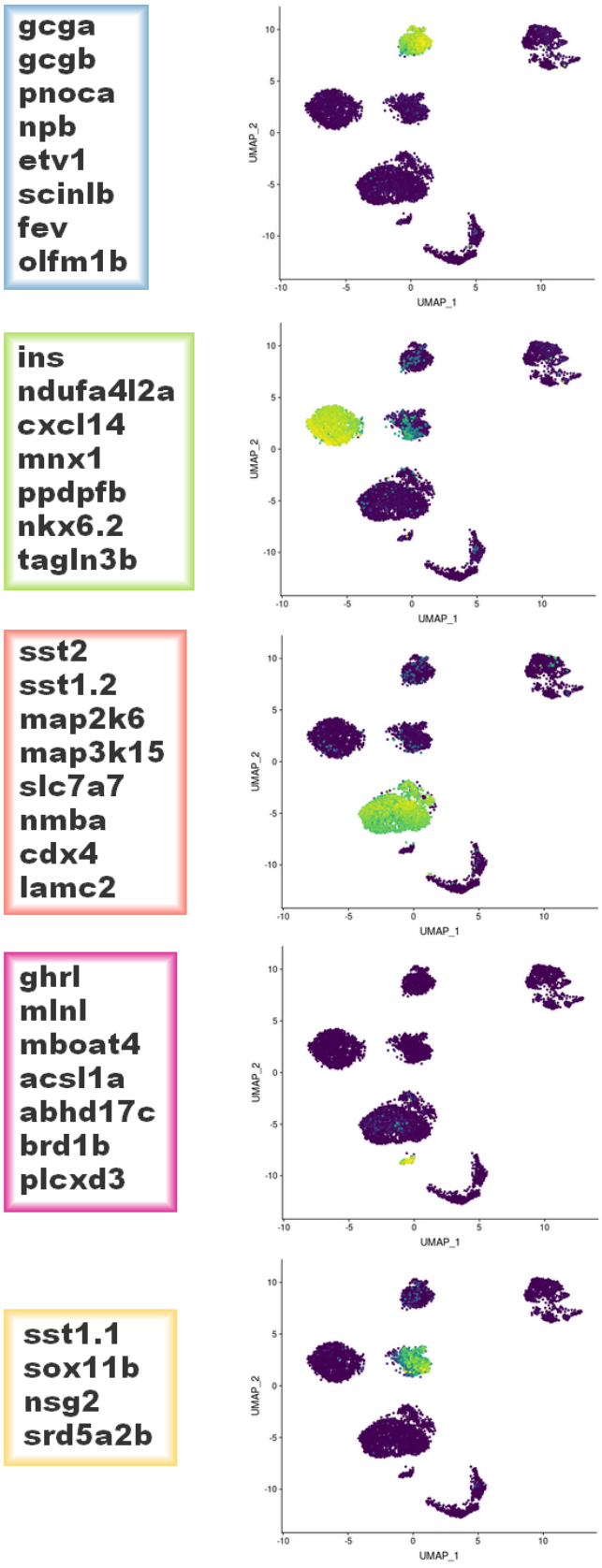


Figure 33: **Identification of marker genes**. The combinations of genes defined for each cell type have been evaluated for their ability to turn on the correct cluster of cells in SC data.

1.4 Characterization of pancreatic cell signatures across human and zebrafish

1.4.1 Human scRNA-Seq from adult pancreatic tissues

The interspecies comparison of the pancreatic signatures that we previously performed indicated that there is a relatively low conservation of the cell-identity signatures for the endocrine cell subtypes [146]. While we were confident with the good quality of the bulk RNA-seq data used for this interspecies analysis, we got several feed-backs from scientists that this low conservation could be due to low quality RNA-seq data (i.e. low purity of cell types). As scRNA-seq data were recently performed for human pancreatic cells, we decided to use these data to verify our previous conclusions. We reanalyzed them with the same methods as for zebrafish in order to be coherent within the two species. In the following sections, I will present key results from separate processing of each dataset before comparing them. Multiple experiments were publicly available from public depositories, and we decided to use three of them to perform cross validation of the signatures obtained in each one.

1.4.1.1 Segerstolpe et al.[136] We retrieved scRNA-Seq data from all 6 healthy individuals used in the study. After cell validation we kept an amount of 1246 cells with on average 5587 (± 2113) genes detected by cell. We first observed how the expression of key markers of different pancreatic population were distributed. We then established scores for each population (see Methods), and we evaluated their pattern within the clusters. By this method, we identified clusters for all 7 major pancreatic cell types with the following cell distribution: 125 acinar cells, 188 ductal cells, 563 alpha cells, 209 beta cells, 64 delta cells, 92 gamma cells and 5 epsilon cells (Figure 34). We obtained on average 16703 (± 3055) genes detected by cluster. We used the same threshold as for zebrafish data to perform differential expression analysis. We identified 39, 46, 8, 9, 10, 293 and 156 genes with significative enriched expression for alpha, beta, delta, gamma, epsilon, ductal and acinar cells respectively and we defined those as marker genes according to our definition (see Methods).

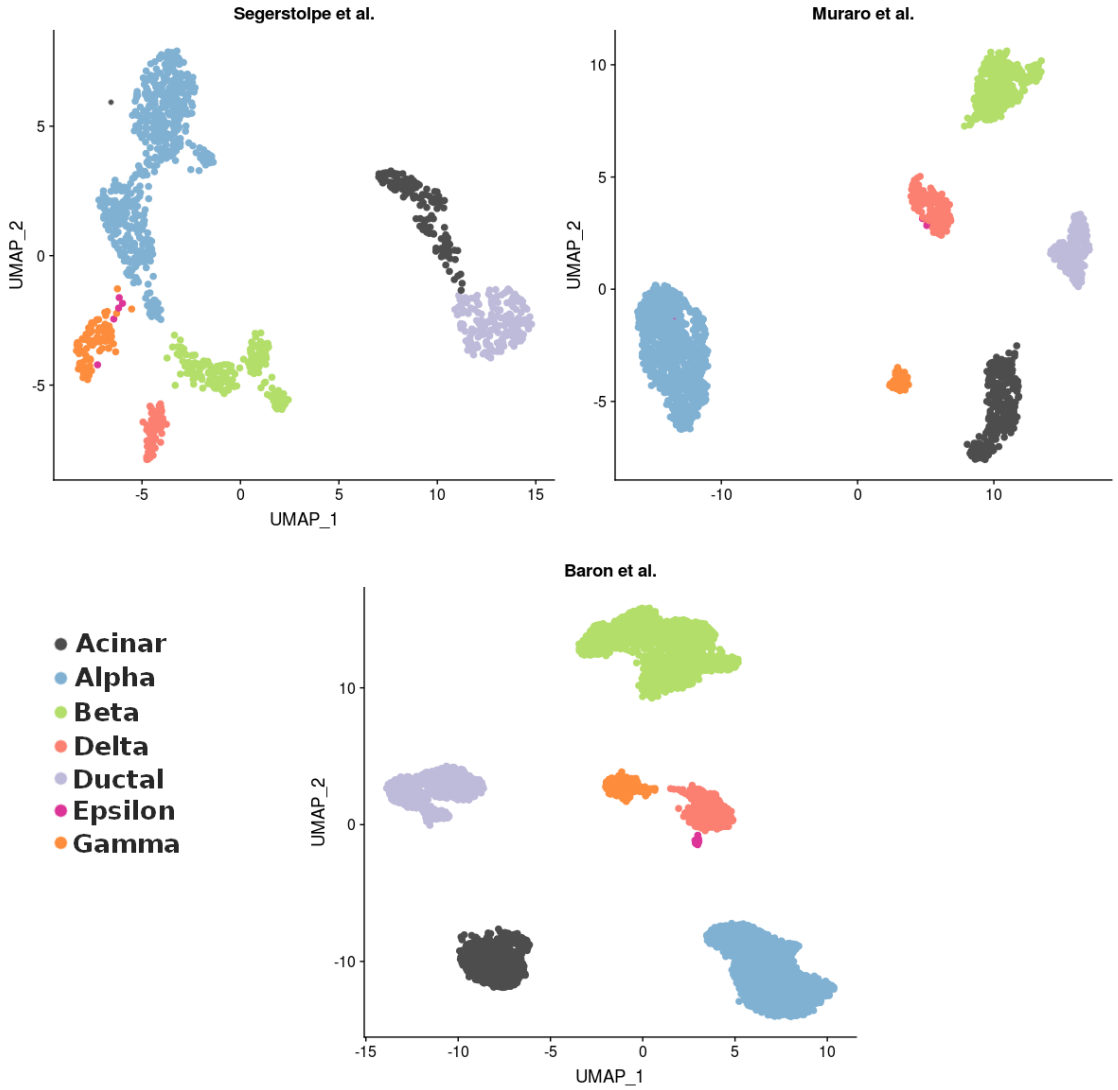


Figure 34: **UMAP embedding of human datasets**. Each dataset is represented by a cell population of 1246 (Segerstolpe), 1815 (Muraro) and 7558 cells (Baron). Cell clusters were labeled according to the key genes they express.

1.4.1.2 Muraro et al.[135]

Data regroup cells from 4 different donors. We validated 1815 cells with on average 5287 (± 1419) genes detected by cell. After clustering, cells were distributed between populations: 283 acinar cells, 187 ductal cells, 778 alpha cells, 335 beta cells, 165 delta cells, 64 gamma cells and 3 epsilon cells. An average of 14674 (± 3011) genes are detected by cluster. We performed differential expression analysis and we identified 35, 130, 29, 14, 4, 369 and 186 marker genes for alpha, beta, delta, gamma, epsilon, ductal and acinar cells, respectively.

1.4.1.3 Baron et al.[134]

We processed data from 4 different donors, and we validated a total of 7558 cells which represents the biggest dataset we retrieved for human adult tissue. The average of genes detected by cell was 1857 (± 596), and after clustering process we obtained the following cell distribution: 977 acinar cells, 908 ductal cells, 2301 alpha cells, 2498 beta cells, 597 delta cells, 260 gamma cells and 17 epsilon cells. The clusters present on average 13368 (± 2328) expressed genes and after performing differential expression analysis, we defined a list of 31, 37, 9, 14, 6, 218 and 110 marker genes for alpha, beta, delta, gamma, epsilon, ductal and acinar cells, respectively.

1.4.2 Cross-validation of Human pancreatic cell markers

The first reason we retrieved data from multiple studies was to evaluate how results can vary depending on the dataset we used. Morever, the different single-cell technologies developed in recent years present some variations. The number of cells they allow to work with or the efficiency of the RNA detection for a single cell are two parameters that can impact the results. For identical costs, the choice of a single-cell technology or another will depend on which of the two parameters you want to promote.
A second aspect to consider is that between the different single-cell experiments we used, they do not present the same amount of cells and we also observed heterogeneity between cell types distribution. This has effect on differential expression analysis as we are dependent on what we compare. Data from Segerstolpe and Muraro groups were produced with similar protocols which allow to detect more RNA molecules per cell but for a lower amount of them. Baron group samples are based on a different approach which promotes higher number of cells over RNA detection per cell.

When comparing the three experiments, our concern was to minimize the impact of these sources of variation (batch effect). We decided first to compare the set of marker genes obtained from the analysis of each dataset separately.

1.4.2.1 Comparing scRNA-Seq results between experiments

We started by performing comparison of the lists of markers from all 3 experiments and we observed that only an average of 6.92 % (± 5.09) of all defined marker genes for a cell type are common between all 3 datasets. This value rises up to 26.67 % (± 12.72) by considering genes found in at least 2 experiments. It highlights that working with an experiment or another can have significative influence on results.

Using all 3 sources, we determined sets of pancreatic markers for each endocrine cell type in human. By starting with beta cells, 9 marker genes were identified by all 3 experiments: MAFA, INS, ADCYAP1, IAPP, G6PC2, SURF4, MEG3, CDKN1C, EIF4A2. The first 5 have already been described to be beta cell marker genes and the 4 others have been presented in the originating papers. We identified 5 markers for alpha cells: PCSK2, COTL1, GPX3, SLC7A2, CAMK2G. Surprisingly, GCG did not pass our threshold for the Muraro group experiment. PCSK2 has been described to be responsible on hormone precurors processing and specifically for the release of glucagon from proglucagon [183, 299]. All 4 others genes have been described to be expressed in pancreas. Interestingly, CAMK2G is a calcium/calmodulin-dependent protein kinase which could play role in glucagon release [300]. For delta cells, we highlighted only 3 marker genes: SST, RBP4 and RGS2. RBP4 has been associated with obesity and type 2 diabetes [301–303]. Finally, only GHRL and PPY have been identified by all 3 datasets as markers of epsilon and gamma cells respectively.

Next, we performed the comparison at the level of genes with an enriched expression (DE genes - Methods 4.5). The average gene overlap between all 3 datasets rise up to 14.69 % (± 7.59) and up to 38.24 % (± 13.09) between at least 2 experiments. We also expected the overlap between gene lists of each experiment to be higher. As results, we obtained 144 genes found to be enriched in beta cells (FIGURE). Among them are PDX1, PCSK1, HOPX, TAGLN3 and TSPAN1. For alpha cells, we identified 139 DE genes and we retrieved notably genes like FEV, ARX, MAFB or GC. We identified 38 delta enriched

genes among which we found HHEX, LEPR, SSTR1 or PKIB. Epsilon cells are represented by 5 genes: BHMT, TM4SF5, GHRL, ASGR1 and VTN. TM4SF5 has been described to be overexpressed in pancreatic cancers. Finally, we obtained 66 genes enriched in gamma cells with genes such as ETV1, ARX, PAX6 or PRKACB.

1.4.2.2 Integration of human single-cell datasets We decided to perform a second approach to extract the information from all 3 datasets. After each datasets were analyzed and annotated separately, we integrated all cells together before doing differential expression analysis on this large combined datasets (see Methods 4.5). The global UMAP embedding we obtained allows to observe a strong colocalization of each cell type cluster between all 3 experiments (Figure 35-A & B). The total of 10619 cells presents the following distribution: 3645 alpha, 3048 beta, 820 delta, 413 gamma, 26 epsilon, 1362 acinar and 1305 ductal cells. We defined a novel list of marker for every cell type and we obtained 23, 39, 9, 18, 11, 114 and 344 marker genes respectively (Figure 35-C + Supplemental table 3). Many of these markers were repordted in at least one of the three studies.

Beta cell markers include all 9 marker genes validated by the previous comparison but also many beta-cell enriched known genes such as human beta marker PDX1. The gene list also contains ERO1B which has be linked to the oxidative protein folding of proinsulin [304]. SLC6A6 encodes taurine and beta-alanine transporter and taurine has been described with potential role in insulin gene expression and glucose homeostasis [305, 306]. PPP1R1A gene encodes protein working as an inhibitor of protein-phosphatase 1 which may be important in hormonal control of glycogen metabolism [307].
We identified the same 5 alpha cell marker genes as presented above but we retrieved also known genes like GCG, MAFB or IRX2 [308, 309]. PDK4 gene encodes a kinase playing a key role in regulation of glucose and fatty acid metabolism and homeostasis [310]. TTR gene is found to be marker of alpha cells and has been associated with endocrine tumors of both pancreas and gut [311]. RGS4 gene has been described as a negative regulator of insulin release [312].
For delta cells, we retrieved SST, RBP4 and RGS2 as previously but also HHEX, NLRP1, PSIP1, DHRS2, SEC11S and PCP4. Notably, HHEX encodes a TF which has been first

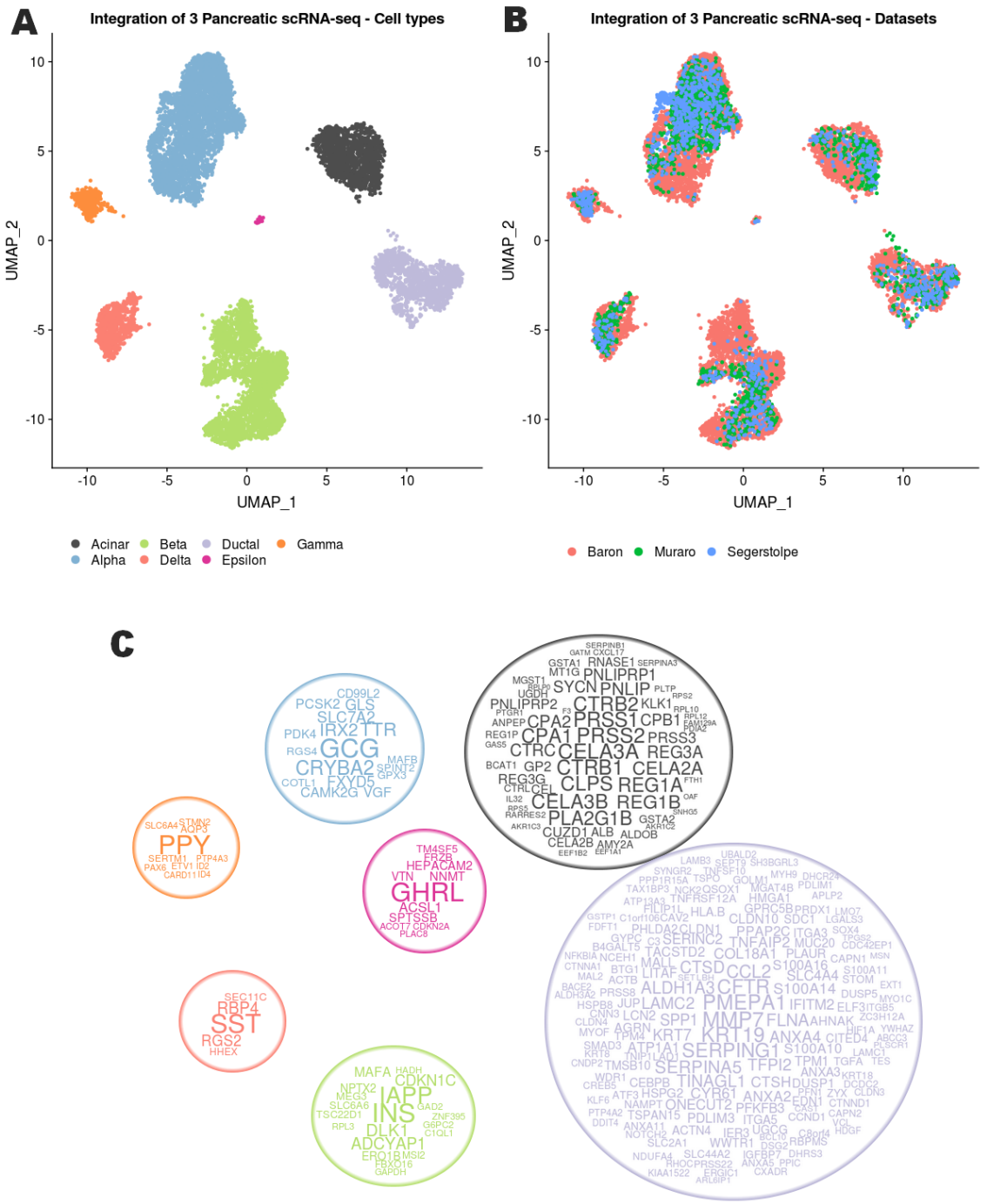


Figure 35: **UMAP embedding of the integrated dataset**. (A) Visualization of the cell type clusters. (B) Visualization of the origin of the dataset. (C) Marker genes identified for each cell type.

described to be linked to T2DM before being identified as specifically expressed in delta cell [185].

We highlighted 11 marker genes for epsilon cells in which we found TM4SF5, ACSL1, SPTSSB, NNMT, VTN, CDKN2A, FRZB, PLAC8, HEPACAM2 and ACOT7 along GHRL. Finally, we identified 18 marker genes for gamma cells. Beside PPY, we found PAX6 and ETV1 as marker genes, but also SERTM1, STMN2, SLC6A4, SCG2, CARD11, ID2, SCGB2A1, PPY, AQP3, C16orf45, DDR1, INPP5F, PTP4A3, ID4, MALAT1 and FXYD2.

1.4.3 Validation of pancreatic endocrine markers across zebrafish and human

An important objective of this work was to define conserved set of genes to be used to identify each endocrine cell type through different species. These sets of genes with conserved expression in a cell type are expected to be responsible on the identity and the function of this cell type.

As we previously showed in our publication, the overlap between marker genes identified in zebrafish and mammals for beta and alpha cells was surprisingly very low, only few enriched genes for one endocrine cell type were conserved between species. Using available scRNA-seq information for human and zebrafish, we decided to perform comparison of signature genes identified in both datasets in order to validate some identity genes across species. If we simply cross the lists of marker genes obtained in each specie for one cell type, the overlap is still very low. For example, only the insulin gene can be highlighted for beta cells in both species. We can mention that expression of genes such as MNX1 and PDX1 showed enrichment in beta cells but as we previously presented *pdx1* is more enriched in delta 1.1 cells in zebrafish and MNX1 in human is not expressed by enough cells (less than 50% of beta cells) to have been considered as beta marker. Meanwhile epsilon identity seems to be highlighted by similar genes in both human and zebrafish. Beside ghrelin expression in both, we found ACSL1 (human) gene and its ortholog *acsl1b* (zebrafish) as interesting epsilon marker, underlying that the acylation of ghrelin has been conserved during the evolution. An other interesting gene is TM4SF5 gene for which ortholog *tm4sf5* did not pass threshold in zebrafish but still showed enriched expression in epsilon cells. If we considered all endocrine cell signals, we observed that expression of hormone genes are conserved as expected in correct cell types between species. The differences in cell population between human and zebrafish make harder the detection of cell signature between them: we identified two somatostatin-expressing cell types in zebrafish while gamma cell type is only present in human. Moreover, these differences led some functions to be remixed and performed by different cell types between species.

We decided to take all sets of identity genes we determined in zebrafish for endocrine cells and we looked at their projection in human map of transcriptomes (Figure 36). We didn't use hormone genes as we wanted to identify new conserved specific genes. Results

showed that alpha signature determined in zebrafish highlighted alpha cluster in human but also gamma cluster. ETV1 and FEV are among genes with enriched expression in human gamma cell. Zebrafish beta signature showed enrichment in human beta cells but also projection into ductal, gamma and delta cells. Surprisingly, delta 1.2 identity didn't present any projection in human delta cells together with slight turn on of ductal cluster. Interestingly, epsilon identity looked like the most conserved of all but low amount of epsilon cells in human datasets limited interpretation of results. Finally, delta 1.1 identity through only SOX11 and SRD5A2 expression did not highlight any particular cluster in human datasets.

Altogether, these results showed that defining a set of markers for each endocrine cell type which display an inter-species conservation is not an easy task. We expected to identify a core of key genes with specific expression in each cell type. However, results suggest that the specificity (or the identity) of these cells could rather be due to differences in the balance of the expression level of key shared genes and this balance could also be different between species.

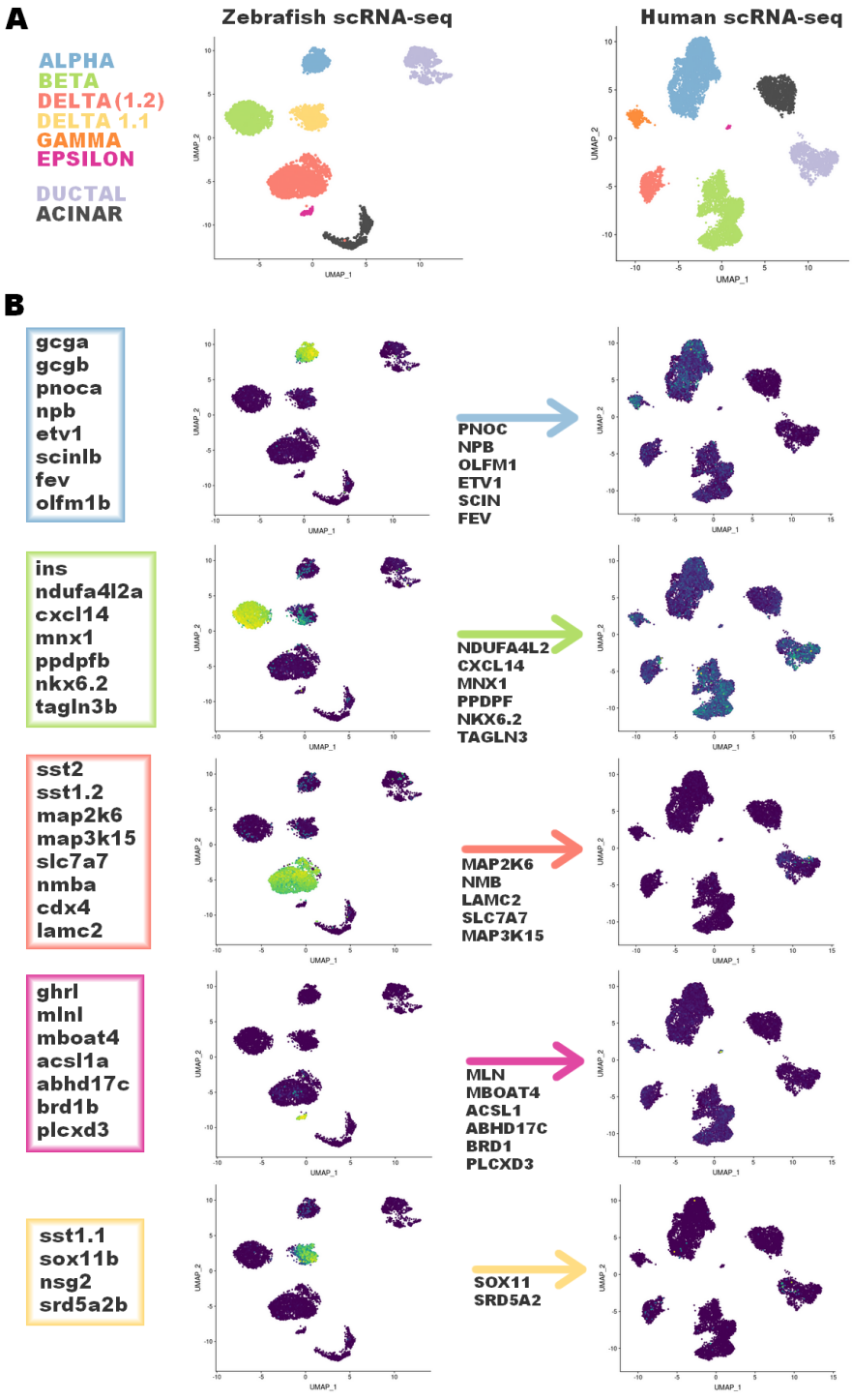


Figure 36: **Projection of the zebrafish endocrine sets of marker genes on human datasets**. Each zebrafish combination has been evaluated for highlighting the expected endocrine cell type in h.

2 Pax6b Gene Regulatory Network during endocrine development in Zebrafish

The second part of this work focused on the transcription factor *Pax6b* in zebrafish and its role in the endocrine differentiation of both pancreatic and enteric tissue. We observed significant similarities between the Pax6-dependent Gene Regulatory Network of both tissues. Notably, the disruption of pax6b in pancreatic endocrine cells (PECs) and enteroendocrine cells (EECs) affects the expression of hormones and the cell fate choice, highlighted by an increase of ghrelin/motilin-like expressing cells in both tissues with a concomitant decrease of other endocrine cell subtypes. PECs and EECs in *pax6b*$^{-/-}$ mainly display tissue-specific transcriptomic changes. However, we highlighted a significative set of genes presenting a similar regulation in both tissues. These results have been combined in a publication presented here, entitled **"Pancreatic and intestinal endocrine cells share common transcriptomic signatures and regulatory programs"**, submitted to BMC Biology (currently under reviewing).

Credit of the work: The study and the characterization of enteroendocrine cells have been conducted by Justine Pirson, Anne-Sophie Reuter and Marianne Voz. They performed the *in situ* hybridization experiments in EECs. Estefania Tarifeño-Saldivia and Justine Pirson carried out the RNA-seq experiments for WT and *pax6b*$^{-/-}$ mutant PECs and EECs. I joined them by performing all bioinformatic analyses concerning the transcriptomes comparison, differential gene expression analysis and cell subtype markers identification. I performed the ISH experiments in PECs and EECs of the *pax6b* mutants.

Pancreatic and intestinal endocrine cells share common transcriptomic signatures and regulatory programs

Arnaud Lavergne †, Estefania Tarifeño-Saldivia #†, Justine Pirson, Anne-Sophie Reuter, Isabelle Manfroid, Marianne L. Voz* and Bernard Peers*.

E-mail : Arnaud Lavergne : arnaud.lavergne@doct.uliege.be; Estefania Tarifeño-Saldivia : etarisal@udec.cl; Anne-Sophie Reuter : as.reuter@alumni.ulg.ac.be ; Isabelle Manfroid : Isabelle.Manfroid@uliege.be; Marianne L. Voz : mvoz@uliege.be; Bernard Peers : bpeers@uliege.be;

Laboratory of Zebrafish Development and Disease Models (ZDDM), GIGA, Avenue de l'Hôpital 1, B34, 4000 SART TILMAN, University of Liège, Belgium

: Present address : Gene Expression and Regulation Laboratory, Department of Biochemistry and Molecular Biology, University of Concepción, Concepción, Chile.

* : co-last authors and corresponding authors

† : co-first authors

Abstract

Background: Endocrine cells of the digestive system, including the pancreatic endocrine cells (PECs) clustered in the islets of Langerhans and the enteroendocrine cells (EECs) scattered in the intestinal epithelium, play an important role in metabolism. Although EECs and PECs are located in distinct organs, they share many features and several common genes control their differentiation. In this study, we investigated comprehensively the similarity of EECs and PECs by defining their transcriptomic landscape and comparing the regulatory programs controlled by *pax6b,* a key player in both EECs and PECs.

Results: RNA-sequencing was performed on EECs and PECs isolated from wild-type and *pax6b* mutant zebrafish. Data mining of wild-type zebrafish EEC data confirmed the expression of orthologs for most known mammalian EEC hormones but also revealed the expression of three additional neuropeptide hormones (Proenkephalin-a, Calcitonin-a and Adcyap1a) not yet reported to be expressed by EECs in any species. Comparison of transcriptomes from EECs, PECs and other zebrafish tissues highlights a very close similarity between EECs and PECs, with more than 70 % of genes being expressed in both endocrine cell types. Comparison of Pax6b-regulated genes in EECs and PECs revealed a significant overlap. *pax6b* loss-of-function does not affect the total number of EECs and PECs but instead disrupts the balance between endocrine cell subtypes, leading to an increase of *ghrelin*- and *motilin-like* expressing cells in both the intestine and pancreas at the expense of other endocrine cells such as beta- and delta-cells in the pancreas and *pyyb*-expressing cells in the intestine. Finally, we show that the homeodomain of Pax6b is dispensable for its action in both EECs and PECs.

Conclusion: This study highlights the close relatedness of EECs and PECs at the transcriptomic and regulatory levels, supporting the hypothesis of a common phylogenetic origin and underscoring the potential implication of EECs in metabolic diseases such as Type 2 diabetes.

Keywords : pancreatic endocrine cells, enteroendocrine, transcriptome, RNA-seq, Pax6, zebrafish, hormone, transcriptome comparison

Background

The enteroendocrine cells (EECs) and pancreatic endocrine cells (PECs) play a crucial role in the control of metabolism of all vertebrates. The PECs, notably via the antagonistic action of glucagon and insulin, secreted respectively by the alpha- and beta-endocrine cells, regulates glycemia. The EECs also participate in this control through the release of some hormones, like glucagon-like peptide-1 (GLP-1) or glucose-dependent insulinotropic polypeptide (GIP), which display incretin effect, enhancing the glucose-stimulated insulin secretion. The EECs secrete more than 15 hormones that modulate many physiological processes including the release of digestive enzymes and bile from the exocrine pancreas and gallbladder, control of intestinal motility, sensing of nutrients and microbial metabolites and stimulation or suppression of appetite [1]. The EECs were initially classified into distinct subtypes according to the hormone they secrete; however, subsequent analyses revealed that there is an extensive co-expression of several hormones in individual EEC [2]. Recent single-cell RNA-seq studies have indeed confirmed this co-expression in murine EECs thereby identifying at least 9 EEC subtypes expressing various combinations of neuropeptides or hormones [3, 4].

EECs display similarities with pancreatic endocrine cells (PECs). Some hormones are produced from both pancreatic and intestinal endocrine cells, like Somatostatin (Sst), Ghrelin (Grhl) or Peptide YY (Pyy) [5, 6]. Also, many transcription factors are involved in the differentiation of both EECs and PECs such as NeuroD, Pax6, Isl1, Nkx2.2, Insm1 and Arx [7–10]. For example, Pax6 is crucial for the differentiation of PECs and its inactivation disturbs the adequate proportion of the different endocrine cell subtypes. Indeed, in *Pax6* KO mice [11, 12] and in *pax6b-/-* zebrafish [13], the number of alpha-, beta- and delta- pancreatic cells is significantly reduced while epsilon-cells are increased. In the small intestine, Pax6 has been reported to be required for the differentiation of some EEC subtypes [14, 15]; however, it is still not known which set of genes are regulated by Pax6 during EECs and PEC differentiation and whether the Pax6-regulated network is shared by these two tissues .

EECs have been detected in the gut of many vertebrate and invertebrate species and their formation appears early during animal evolution dating back to the cnidarian-bilaterian ancestor [16–18]. In contrast, the pancreas and the PECs appeared much later with the first vertebrate species [19, 20]. The similarity between EECs and PECs could be due to several reasons. A first possibility could be a common phylogenetic origin; the PECs could have indeed derived from some EECs and moved out from the gut to form clusters and to secrete other hormones like insulin and glucagon. This scenario is compatible with the "sister-cell-type model" [21] where two cell types derived from an ancestral cell type by segregation of function (e.g. secretion of distinct hormones) due to the gain or loss of transcription factors. In this model, the two sister cell types still share many characteristics and can display similarities in their transcriptomes [22]. Another reason of the PECs and EECs similarities could be the co-option of EEC (or neuronal) regulatory pathways by the pancreatic cells [23]. To approach this matter, we have measured in the present study the degree of similarity between EECs and PECs at the genome-wide expression level and determined whether they share regulatory programs. We previously determined the transcriptomic landscape of zebrafish PECs [24]. In contrast, the EECs from zebrafish have not yet been fully characterized and only a few hormones have been shown be expressed in these cells [25–29]. The first objective of this study was to perform a detailed characterization of zebrafish EECs by determining their transcriptomic profiles and identifying all neuropeptide hormones and regulatory genes expressed in these cells. Secondly, the transcriptomic landscape of zebrafish EECs was compared with those of zebrafish PECs to measure the degree of similarity. Finally, we analysed the consequences of *pax6b* inactivation on the transcriptome of both pancreatic and intestinal endocrine cells and show that Pax6b controls a large set of genes identical in both EECs and PECs further supporting their close relatedness. Finally, we discuss our findings in light of evolution of EECs and pancreatic cells, supporting the model of a common phylogenetic origin, and bringing some implications for deciphering the defects involved in type 2 diabetes.

Results

Characterization of the enteroendocrine cells in zebrafish

In mice, at least 15 different hormones have been described to be expressed in the gut [1, 30]. Out of these, we identified in the zebrafish genome at least one orthologue for 11 of them, 9 showing syntenic conservation at their genomic loci with the murine/human genes, thereby validating orthology relationship (Table 1). We did not find any orthologous genes coding for gastrin (gast), secretin (sct), the pancreatic polypeptide (pp) nor motilin (mln); however, a *motilin-lik*e (*mlnl*) gene, proposed to be the functional equivalent of motilin, is present in the zebrafish genome [31]. In order to identify all peptide hormones expressed in the gut of zebrafish larva, we performed RNA-seq on isolated EECs. This was achieved by microdissection of intestines from 4dpf zebrafish transgenic *Tg(pax6b:GFP)*^{ulg515tg} larvae, which express GFP in the EECs [32], followed by cell dissociation and selection of GFP+ EECs by FACS. The restricted expression of Pax6b:GFP in the EEC and not in the enteric neurons (EN) was confirmed by immunohistochemistry showing a total absence of colocalisation between the Hu enteric marker and pax6b:GFP (Fig S1). RNA-seq was performed on four independent EEC preparations and about 40 million paired-end reads were obtained from each Illumina library, 70-80% of which mapped to the zebrafish genome. Table S1 gives the expression level of all genes in CPM (counts per million) and in FPKM (fragments per kilobase of exon model per million reads mapped). This expression profiling allowed us to determine that, amongst the 12 zebrafish genes corresponding to EEC hormones, 11 are expressed in the zebrafish EECs at significant levels (above 100 CPM or 100 FPKM), *neurotensin* (*nts*) being the only hormone not expressed in EECs (Table 1). The expression of many of these hormones was further validated by whole-mount *in situ* hybridization (WISH) thereby locating the region of the gut displaying strongest expression and defining the onset of their expression (Fig.1 and Table1). The number of EECs detected in the gut of 4dpf zebrafish larvae by WISH varies for each hormone; for example, more than 20 cells expressing *peptide YY-b* (*pyyb)*, *cholecystokinin-a* (*ccka)*, *preproglucagon-a* (*gcga)*, *galanin* (*galn) or insulin-like 5a (insl5a)* transcripts were detected (Fig. 1A,B,G,H,I; Fig.S2 and Table1) while only few cells express

ghrelin (ghrl), *glucose-dependent insulinotropic polypeptide* (*gip)* and peptide YY-a (pyya) (Fig. 1D-F and Table1). These differences of expression are also reflected in the RNA-seq data : *gcga*, *ppyb*, *ccka*, and *insl5a* reaching more than 2000 CPM while *ghrl* and *gip* range around 100 CPM (Table 1).

Data mining of the zebrafish EEC RNA-seq with the GO term "hormone" also revealed a strong expression of three neuropeptide transcripts not described so far as expressed in mammalian EECs: *adenylate cyclase-activating polypeptide 1a (adcyap1a*)(previously named *PACAP*), *proenkephalin-a* (*penka*) and *calcitonin (calca*) (Table 1). Fluorescent *in situ* hybridization experiments (FISH) confirmed the expression of these transcripts in the zebrafish EECs (Fig. 1 J-L). An average of 68 *penka*+ and 12 *calcitonin*+ EECs were detected in the bulb and the mid-gut and more than 80 cells expressing *adcyap1a* were found all over the gut (Table1 and Fig.S2). As enteric neurons (ENs) surround the gut at that stage and express a variety of neuropeptides, some being expressed also by EEC, we next determined whether some of the hormone expressed by EEC would be also detected in ENs. For that purpose, we retrieve the transcriptomic data of zebrafish ENs, recently determined by RNA-seq after sorting the ENs by FACS from *Tg(phox2b::GFP)* larvae at 7 dpf [33]. Comparison of the ENs and EECs RNA-seq data indicate that most EECs hormones display a much higher expression in EECs compared to ENs, except for *sst1.2*, *sst2* and *vipb* (see Table 1). FISH using the EN specific marker *Phox2b* confirms the strong expression all over the gut of *vipb* in ENs while the *vipb*+ EECs are mostly located in the posterior gut (Fig. S3A and B). Interestingly, RNAseq data indicate that adcyap1a is exclusively expressed by EECs while adcyap1b is restricted to ENs. FISH confirms the strong expression of adcyap1a in EECs and not in *phox2b*+ ENs (Fig S3 C). However, IHC using an anti-ADCYAP1 antibody reveals staining in pax6b:GFP+ EECs as well as at the level of EN axon fibers in larvae, as previously shown [34] (Fig.2B). Similar results were observed in the adult intestine (Fig. 2D). Taken together, all these data confirm the expression of the Adcyap1a in EECs and of Adcyap1b in ENs. As for the two other new hormones Calca and Penka, RNAseq comparison indicate that they are exclusively expressed

by EECs and this was confirmed for Penka for which we observe strong Enkephalin immunostainings in EECs but not in ENs at larva stage as well as in adults (Fig. 2 A and C).

Close relatedness between the enteroendocrine and pancreatic endocrine cells.

Although EECs and PECs are known to share common features, we wanted to evaluate the degree of similarity at a more global level. Thus, we compared the RNA-seq data of the zebrafish PECs and EECs as well as of other zebrafish cell types and organs used as control (Fig 3A). The heatmap shows that all endocrine cells cluster together apart from other pancreatic or intestinal cells and other tissues. Indeed, the EEC cluster is much closer to the PEC cluster than to the ductal or the acinar pancreatic cells or even to the whole intestine tissue which is mainly composed of enterocytes and of only 1-2% of EECs. In agreement with previous reports showing shared features between pancreatic cells and neurons [23, 35], this clustering analysis reveals also transcriptomic similarities of brain tissue not only with pancreatic but also with intestinal endocrine cells.

The close relatedness of EECs and PECs is also illustrated by the high percentage of genes expressed in both tissues: among all genes expressed in pancreatic or intestinal endocrine cells above the threshold of 10 normalized CPM (8326 and 8264 genes, respectively), 74% of them are expressed in both tissues (Fig. 3B), while if the same analysis is done for genes expressed in EECs and liver, only 45% of them are expressed in both tissues (data not shown). Similar percentages are obtained when considering genes coding for transcription factors (TF): 72% of all expressed TF-coding genes are detected in both PECs and EECs, while only 45% are expressed in both pancreatic endocrine cells and liver. The list of 481 TF expressed in both PECs and EECs (given in Table S2) includes notably most of the transcription factors reported to be important for zebrafish PEC differentiation (Table 2). For example, *arxa* and *pax4*, known to be important determinant of cell fate of pancreatic endocrine cells [36, 37], are also expressed in the zebrafish gut and, like in pancreas, their expression is mainly non-overlapping (Fig 4A). *Fev*, known to be highly expressed in pancreatic endocrine cells [38] is also among of the most highly expressed transcription factors in the zebrafish

EECs (Table 2) and is detected by FISH in many cells all over the gut (Fig. 4B). Similarly, *insm1b* [9] is also highly expressed in zebrafish EECs (Fig. 4 C-D and table 2) in addition to its expression in enteric neurons (labelled with asterisk in Fig. 4D) as described in mouse. Expression of *pdx1*, visualised using the *tg(BAC pdx1:GFP)* [39], is also detected in scattered EECs of the mid- and posterior-gut (Fig 4E and G), in addition to a widespread expression in cells of the rostral part of the intestine as previously described [40] (Fig. 4 E and F). The expression of *pdx1* in EECs of the mid and posterior gut was confirmed by WISH (figure 4H) and by immunohistochemistry (Fig. 4 I) As expected, *ascl1a* was barely detected in the transcriptome of 4 dpf EECs (Table 2) as we have previously shown that its expression in the secretory precursors cells turns off as soon as the cells pursue their differentiation process [27]. In the same way, *sox4b*, coexpressed with *ascl1a* in the secretory precursor cells, is hardly detected in the transcriptome of the EECs at 4 dpf (Table 2), indicating that *sox4b* is also switched off with the maturation of the EECs. In contrast, Nkx6.1, Nkx6.2 and Mnx1 are the only 3 transcription factors involved in zebrafish PEC differentiation that are not expressed in EECs. Many of these EEC/PEC transcription factors are expressed at lower levels in enteric neurons as revealed by the ENs RNA-seq data (Table 2) with the exception of insm1b, neurod1 and ascl1a which are expressed above 50 CPM.

The number of genes which are selectively expressed in PECs or in EECs is relatively low (gene lists available in Table S3). Gene ontology (GO) enrichment analyses of the PEC-specific genes highlight terms such as “cell adhesion” due to the enriched expression of many cadherins and other cell adhesion molecules. This enrichment makes senses as the PECs are clustered into islets while EECs are scattered in the gut. The EECs-specific genes display an enrichment of GO terms including “drug transmembrane transport”, “ABC transporter” or “oxidation process” due to the enriched expression of many transmembrane transporters and cytochrome P450 type molecules (Table S3).

In conclusion, all these data indicate a close relatedness between PECs and EECs at the level of their transcriptomes and of their regulatory factors, suggesting the involvement of similar regulatory programs in these two cell types.

Shared Pax6-dependent regulatory programs between pancreatic and intestinal endocrine cells.

To investigate whether pancreatic and intestinal endocrine cells are controlled by similar regulatory programs, we analysed the effect of Pax6 inactivation on the transcriptome of both cell types. Zebrafish has two *pax6* paralogs; however, only *pax6b* is expressed in PECs and EECs. To isolate *pax6b* mutant EECs and PECs, *pax6b*^{sa0086} heterozygous fish harbouring the *Tg(pax6b:GFP)* transgene were inbred and *pax6b*^{sa0086} homozygous embryos were selected based on lens abnormalities [13]. The expression of the transgene *Tg(pax6b:GFP)* was not perturbed in the homozygous mutants (data not shown), allowing us to isolate by FACS the PECs or EECs from *pax6b-/-* mutant larvae after microdissection of the dorsal pancreatic bud or intestine, respectively. RNA-seq was performed on three independent preparations of EECs or PECs purified from *pax6b*^{sa0086} mutant and from wild-type zebrafish (see Materials and Methods). Principal component analysis shows a tight clustering of replicates underscoring a good reproducibility of the data and showing distinct clusters for mutant and wild-type cells for both PECs and EECs (Fig. 5A). Differential expression analyses identified 2824 and 1634 Pax6b-regulated genes in PECs and in EECs, respectively, (FDR<0.1)(Fig. 5 B,D)(expression of all genes and of Pax6b-regulated genes are given in Table S4 and S5, respectively). Comparison of these two sets of genes reveals that there is a large set of 517 genes which are altered by *pax6b* inactivation in both organs (Fig. 5C); this overlap is statistically significant according to the Fisher's Exact test (p-value < 2.2e-16). Furthermore, most of these genes (424 genes) are regulated in the same way in the pancreas and in the intestine (215 up-regulated and 209 down-regulated genes)(Fig. 5C), indicating that a part of the Pax6-dependent gene regulatory network is thus identical in EECs and PECs. To explore the pathways and biological processes regulated by Pax6b, we performed a gene ontology analysis on these 3 sets of Pax6b-regulated genes : the common set for PECs and EECs (517 genes) or the specific one for PEC (2307 genes) or EECs (1117 genes)(see Table S5). GO terms associated to "cell-cell signalling", "regulation of appetite" or "hormone activity"

were enriched significantly in the common set as several hormones expressed in PECs and EECs are affected in the pax6 mutant (e.g. *pyya*, *sst2, pyyb*,… ; see below). Such an effect results probably from a change in the expression of several transcription factors (e.g. *pdx1, neurod1, rfx6, nkx2.2*, …; see below) as revealed in the GO lists "endocrine pancreas development" and "sequence-specific DNA binding". Also, the GO analysis highlights the role of Pax6b in the control of ion transport (K+, Ca++ and Cl-) as well as calcium signalling. Interestingly, "ion transport" and "calcium channel / binding activity" were also GO enriched terms among the genes regulated by Pax6b only in PECs or in EECs, supporting an important role of Pax6b for controlling these pathways.

Drastic increase of ghrelin/motilin-like expressing cells in pax6b$^{-/-}$ pancreatic and intestinal endocrine cells with a concomitant decrease of other endocrine cell subtypes.

The above analyses indicate a regulation by Pax6 of the expression of several hormones in PEC and EEC. Among them, *ghrelin* (*ghrl*) and *motilin-like (mlnl) genes* are strongly up-regulated in both the pancreas and intestine of *pax6b* mutants (Fig. 5 E and F, Table 3). FISH experiments confirmed an increase of the number of *ghrl*+ cells and *mlnl*+ cells in both pancreas and intestine (Fig. 6 A-H). Furthermore, double fluorescent staining demonstrated that these two hormones are often co-expressed in the same endocrine cells (Fig. 6 I and J). In both tissues, the increase of *ghrl* and *mlnl* gene expression is concomitant to a decrease of other hormones : *galn, pyyb* and *calca* genes are indeed significantly down-regulated in EECs of *pax6b^{sa0086}* mutants, while in the pancreas, expression of notably the *insulin* (*ins*) and *somatostatin-1.1, -1.2 and 2* genes is strongly decreased (Fig. 5E and Table 3). To determine whether these modifications in hormone expression result from a change in cell subtype proportion in *pax6b* mutant PECs, we analysed the expression of markers of these different pancreatic cell subtypes that we and others have previously identified (Table S6) [24, 41]. We found that about 40% of genes showing an enriched-expression in beta-cell (with a fold enrichment > 4-fold) were down-regulated in the islet of *pax6b-/-* embryos, while only 2% were up-regulated (Fig 7 upper panel), strongly suggesting a loss of beta-cells. This loss was

further confirmed by FISH experiments: firstly, *ndufa4l2,* a gene showing a restricted expression in zebrafish pancreatic beta-cells, was not detected anymore in the *pax6b* mutants (Fig. 7A-D). Similarly, expression of the two beta-cell enriched transcription factors *pdx1* and *nkx6.2* were strongly decreased in PECs of *pax6b* mutants while could still be detected at low levels in the ventral pancreatic progenitors (Fig. 7E-H, Table 3). All these data indicate the loss of beta-cells in *pax6b*-/-^{sa0086} embryos. Concerning the pancreatic delta-cells, 35% of delta-enriched genes were down-regulated in the *pax6b*^{sa0086} mutants and 12% were up-regulated (Fig 7 upper panel). The decreased number of delta-cells in *pax6b* mutants was further supported by FISH through the strong reduction of cells expressing *somatostatin2* (*sst2*) as well as the loss of the delta-specific markers *laminin C2* (*lamc2*) and the hormone peptide *Neuromedin Ba* (*nmba*)(Fig 7 I-L, Table 3). The two delta-specific transcription factors Cdx4 and Hhex were also down-regulated in PECs of *pax6b* mutants (Table 3). Inversely, 40% of genes with enriched-expression in pancreatic epsilon-cells are up-regulated in *pax6b* mutants while only 2% are decreased (Fig. 7 upper panel). Like *ghrelin* and *motilin-like*, *mboat4,* another epsilon-specific marker is strongly stimulated following *pax6b* inactivation as revealed by FISH (Fig 7 M,N) and by the RNA-seq data (Table 3). Thus, these data also indicate an increase of epsilon-cells in *pax6b*^{sa0086} mutant embryos. Finally, the alpha-cells in pax6b mutants display an “ambiguous” phenotype with a respective decrease and increase of 26% and 12% of alpha-enriched genes (Fig 7 upper panel). The RNA-seq data indeed reveal a strong decrease of the alpha-cell markers *gcgb* and *pnoca* while others like *gcga*, *arx* and *scinlb* are even increased (Table 3). FISH confirmed these results. Indeed, although *gcga* and *gcgb* are detected in the same alpha-cells in wild-type embryos, *gcgb* expression was severely reduced in the *pax6b* mutants in contrast to *gcga* that remained highly expressed (Fig 7 I and J). Like *gcgb*, the alpha cell marker *pnoca* (*prepronociceptin a)* was drastically down-regulated (Fig 7 Q, R). On the other hand, the alpha cell marker *scindl* (Fig 7 U,V) and the alpha-cell determinant *arxa* (Fig 7 S,T) were still expressed. Double labelling of *arxa* and *scinlb* with *ghrelin* revealed that these two markers are actually expressed in both alpha and epsilon cells in zebrafish (Fig 7 S-X). Regarding the pan-endocrine markers, which are not enriched in a

specific endocrine cell subtype, only a minority (about 15%) are either down- or up-regulated in the pax6b$^{-/-}$ mutants (Fig 7 upper panel). Altogether, our data indicate that there is a loss of beta-cells and a strong reduction of delta-cells in the *pax6b*$^{-/-}$ mutants with a concomitant increase in the number of epsilon-cells. Alpha-cells are still generated in the *pax6b*^{sa0086} mutants but these cells are misdifferentiated.

To verify that the increase of *ghrl+/mlnl+* endocrine cells in the intestine of *pax6b*$^{-/-}$ larvae (Fig. 6) is also concomitant to a decrease of other EECs, we analysed by FISH the expression of *pyyb*, a gene highly expressed in the intestine and strongly down-regulated in *pax6b* mutants (Table 3). The number of EECs expressing *pyyb* was indeed drastically reduced in the intestine of *pax6b*^{sa0086} mutants (Fig 8 D and E). Thus, all these FISH confirms that *pax6b* has a similar role in the control of the proportion of the various endocrine cell types in both intestine and pancreas.

Like in the pancreas, the homeodomain of Pax6b is not crucial for the differentiation of the endocrine cells of the intestine.

Previous studies have shown that a missense mutation in the homeobox of *pax6b* (sunrise allele) gene causes eye developmental defects but does not perturb the development of pancreatic endocrine cells [13, 42], indicating that the homeodomain of Pax6b is not crucial for endocrine cell differentiation in pancreas in contrast to eye development. To get a clue whether the mode of action of Pax6b is also similar between EECs and PECs, we next investigated whether the EEC differentiation was affected in the sunrise mutants. FISH revealed no effect of the sunrise mutation on the expression of *pyyb* and of *mlnl* in contrast to the null *pax6b*^{sa0086} mutation (Fig 8 C and F). This indicates that the homeodomain of Pax6b is not crucial for the differentiation of endocrine cells both in intestine and in pancreas.

Discussion

By RNA sequencing of FACS-purified cells, we report in this study the transcriptome of enteroendocrine cells (EECs) from zebrafish highlighting the repertoire of peptide hormones and regulatory factors expressed by these cells. The comparison of these RNA-seq data with other zebrafish tissues confirms the close similarity of EECs with pancreatic endocrine cells (PECs). Furthermore, we show that Pax6b is required for the proper expression of many hormones in both the pancreas and intestine, acting via a large set of common gene targets. Inactivation of zebrafish Pax6b does not affect the total number of endocrine cells but affects the balance between endocrine cell subtypes, leading to an increase of ghrelin+/motilin-like+ cells and a reduction of other endocrine cell subtypes in both tissues. The Pax6b homeodomain is also dispensable in the two organs. All these data highlight the high similarity between EECs and PECs.

The present study reveals also that the zebrafish EECs express most of the hormones known to be secreted by mammalian EECs. Only 4 EEC hormones were not detected in zebrafish. Secretin (Sct), Pancreatic polypeptide (PP) and Gastrin could not be detected as no orthologous gene is present in the zebrafish genome. The absence of *sct* and *pp* genes is in agreement with previous studies indicating the lack of these two genes in teleost [43–45]. The detection of PP by immunofluorescence at the level of zebrafish EECs [29] and zebrafish PECs [46] is probably due to cross-reactions of the antibodies with Pyya and/or Pyyb as significant sequence similarities is observed between Pyy and PP mature peptides (see Fig. S4). Although the *neurotensin* gene is present in the zebrafish genome, no expression was detected in EECs, at least in 4 dpf larvae. On the other hand, the present transcriptomic profiling of zebrafish EECs reveals the identification of three neuropeptide transcripts not described so far in EECs : *proenkephalin-a* (*penka*), *calcitonin* (*calca*) and *adenylate cyclase activating polyeptide 1a* (*adcyap1a*). We could confirm at the protein level the expression of two of them, Enkephalin and Adcyap1 (figure 2). Further studies should be performed to confirm the expression of Calca at the protein level and determine the role of these 3

neuropeptides on the digestive tract. To get a hint whether these hormones could be also expressed in the mammalian intestine, we searched into the NCBI GEO repository and into the human protein atlas database (https://www.proteinatlas.org/humanproteome/tissue)[47] for the presence of these transcripts in the gut. We found that *Penk* is expressed in the colon of 8 week-old male mice [48] and at low level in the human colon. Similarly, human *CALCA* and more specially *CALCB* (Calcitonin related polypeptide B) transcripts can be detected in bulk RNA-seq of the gastrointestinal tract mainly at the level of the colon. Finally, ADCYAP1 expression is detected in the human gastrointestinal tract mostly in the colon as well as in beta pancreatic cells [49]. Furthermore, the recent single-cell RNAseq data from the small murine intestine reveal that some enterochromaffin cells express *Adcyap1* at low levels [3]. Thus, all these data suggest that the three hormones Penk, Calca(/b) and Adcyap1 could also be expressed in some mammalian EECs and their expression should be verified by immunostaining notably at the level of the colon.

Some similarity between EECs and PECs has been previously recognized notably thanks to the shared expression of markers including some hormones and transcription factors. Using the zebrafish, the present study demonstrates the similitude at the transcriptomic level and by identifying a common Pax6b-dependant gene regulatory network. When considering the ontogeny and evolutionary aspects of EECs and PECs, it is tempting to speculate that these two endocrine cell types not only share developmental pathways but also have a common ancestor. Indeed, on the ontogeny aspect, both EECs and PECs derive from the endoderm layer and their specification is controlled by lateral inhibition via the Delta/Notch signaling pathway, leading to the formation of scattered endocrine cells within the epithelium. Inhibition of the Notch pathway allows the activation of proneuronal bHLH factors in endocrine precursors that drive the formation of EECs and PECs [50, 51]. As proposed by others [52], the EECs have probably co-opted for these neuronal gene programs to trigger their differentiation early during evolution of metazoa. While the EECs are present in the gut of invertebrates, the first signs of the pancreatic organ appears much later with the first vertebrates : the agnathan species, such as lampreys or hagfishes, possess at the adult stage

one or two aggregate(s) of insulin-expressing cells just next to the intestine and near the bile duct without associated exocrine tissue while other endocrine cells expressing glucagon, somatostatin or PYY-like peptides are present within the intestinal epithelium [19, 20, 53]. Furthermore, at larva stage, some of these insulin+ cells are still detected within the intestinal epithelium of lampreys [54]. Thus, these observations suggest that, when the first vertebrate species appeared, some enteroendocrine cells could move out from the gut to form clusters, possibly through the up-regulation of cell adhesion molecules (Table S3), that have led to the formation of pancreatic endocrine islets. This hypothesis is supported by studies on sea urchin larvae showing the presence of cells expressing insulin-like peptide in the gut region located at the stomach-intestine boundary and expressing the pdx1 orthologous gene [18, 55]. Furthermore, the generation of the first PECs in the zebrafish embryos is also consistent with this idea : indeed, the first insulin expressing cells appear scattered within the anterior endodermal gut epithelium at an early stage (14 hpf) and subsequently migrate dorsally from the embryonic gut to form a cluster and generate the principal pancreatic islet (named the dorsal pancreatic bud)[46]. Thus, all these observations entail the notion that PECs are in fact EEC-like cells, explaining why PECs share so many features with EECs. This concept is also supported by previous reports showing the conversion of EECs into insulin-expressing beta-like cells upon targeted expression or inactivation of some transcription factors [56–58]. The close relatedness between EECs and PECs underlined in our study has also important implications for the interpretation of GWAS for type 2 diabetes (T2D). Indeed, as many of the genomic variants linked to T2D are near (or in) genes expressed in PECs [59, 60], it is widely considered that at least part of the defects causing T2D stems from PEC dysfunctions. However, as the transcriptomes of PECs and EECs are highly similar, the defects causing T2D may also have their origin from EECs.

The present study confirms that Pax6 is required for the proper expression of many hormones from EECs and PECs and change the proportion of the endocrine cell subtype without affecting the total number of endocrine cells in both the pancreas and intestine. The pancreatic phenotype of zebrafish *pax6b* null mutants is comparable to the *Pax6* mutant mice

as the same major pancreatic hormones are up- and down-regulated in the two species, although the effect on beta-cells looks more drastic in zebrafish than in mice (Pax6 KO mice displaying only 70% reduction)[11, 61]. The almost complete loss of beta-cells in zebrafish may derive from a stronger decrease in the expression of the beta-specific factor Pdx1 (as well as Mnx1 and Nkx6.2) in zebrafish pax6b mutants. On the other hand, in the intestine, there is a striking contrast between zebrafish and mice Pax6 mutants : while Pax6 is required for the expression of several EEC hormones in both species, the identity of these hormones is strikingly divergent. Indeed, among the 8 EEC hormones analyzed in Pax6 mutant mice, *Gcg* (GLP1 and 2), *Sst* and *Gip* genes were strongly downregulated while *Pyy* was not significantly affected (Ghrelin and Motilin were not analyzed in Pax6 KO murine intestine). In zebrafish, *pax6b* inactivation causes a drastic down-regulation of *pyyb* and does not affect significantly *gcga*, *sst2* and *gip* gene expression. The reason for such differences between mice and zebrafish is unclear and probably rely on changes during vertebrate evolution either at the level of genomic Pax6 binding sites or of downstream regulatory programs. Further studies including ChIP-seq experiments will be required to answer that question.

Methods

Zebrafish maintenance, transgenic and mutant lines, isolation of EECs and PECs by FACS.

Zebrafish (Danio *rerio*) were raised according to standard protocols and staged according to Kimmel [62]. The following zebrafish transgenic and mutant lines were used : *Tg(pax6b:GFP)*^{ulg515} [63], *TgBAC(pdx1:EGFP)*^{bns13} [39] , *Tg(-8.5nkx2.2a:GFP)*^{ia2} [64], *pax6b*^{sa0086} and *pax6b*^{sunrise} [13]. Enteroendocrine cells (EECs) were isolated by dissecting the gut from about 200 *Tg(pax6b:GFP)*^{ulg515} larvae at 4 dpf, taking care of not including pancreatic tissue. Cell dissociation was next performed by incubation in HBSS 1x supplemented with 100 U/ml collagenase IV and 0.3 U/ml Dispase (Life Technologies) for 10 minutes. Cells were washed in HBSS (Mg^{2+} and Ca^{2+} free) containing 1% BSA and GFP-expressing EECs were selected by two consecutive FACS purifications, the first in the “yield

mode" and the second in "the purity mode", using FACS Aria II. Four replicates of EEC containing about 3000 cells were prepared. Pancreatic endocrine cells (PECs) were also obtained from the *Tg(pax6b:GFP)*^{ulg515} line [63] by dissecting the dorsal pancreatic bud from about 200 27-hpf transgenic embryos. FACS selection was performed as described for EECs except that cell dissociation was performed in *Tryple Select* 1X (Gibco) supplemented with 100 U/ml collagenase IV (Life Technologies) for 5 minutes. For the preparations of EECs and PECs from *pax6b* null mutant embryos, the *pax6*^{sa0086} line [13] was first crossed with the *Tg(pax6b:GFP)*^{ulg515} line; heterozygous *pax6b*^{sa0086} fish harboring the transgene (pax6b:GFP) were inbred to generate homozygous *pax6b*^{sa0086} transgenic embryos which were selected based on the absence or reduction of lens. The isolation of EECs and PECs from *pax6b*^{sa0086} homozygotes were performed in triplicates following the same procedure than for the wild-type larvae. The accuracy of *pax6b*^{sa0086} homozygotes selection was verified after the RNA-seq by checking the presence of the null sa0086 allele in 100% of *pax6b* reads in the mutant samples.

cDNA synthesis, library preparation and sequencing.

Each EEC or PEC sample obtained after FACS was directly pelleted by centrifugation and resuspended in 3.5 µl of *reaction buffer*, lysed by freezing in liquid nitrogen and stored at -80°C according the the Smart-seq2 protocol [65]. cDNA was synthesised and amplified by a 13 cycles PCR reaction. Quality of cDNA was verified by 2100 High Sensitivity DNA assay (Agilent technologies). 1 ng cDNA was used for preparing each cDNA library using Nextera-XT kit (Illumina) and sequenced on Hi-seq 2000 to obtain around 40-60 millions of reads (100 base paired-ends).

RNA-seq data analyses

Sequences were trimmed in order to remove adaptors and low quality bases. Trimmed reads were mapped in to the zebrafish genome GRCz11 (Ensembl Release 92, www.ensembl.org) using STAR software v.2.5.4b [66]. Gene expression was measured from the mapped reads by using built-in STAR module (--quantMode GeneCounts) and are expressed in counts of

reads per million (CPM) [67]. The RNA-seq raw data have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE149081. The comparison of the transcriptome of EECs and PECs with other zebrafish tissues was performed using DESeq2 R package [68]. This comparison includes RNA-seq data from the heart (ArrayExpress: E-MTAB-460; GEO: GSE71755), brain (ArrayExpress: E-MTAB-460), pancreatic cells [24, 69] , liver (GEO: GSE82246) and intestine (GEO: GSE83195). The reproducibility of RNA-seq data from wild-type and *pax6b* mutants was verified by a principal component analysis (PCA) obtained using the DESeq R package [70]. Differential expression (DE) analysis was performed using the R package DESeq2 [68], to identify all genes displaying significant change in expression between wild-type and *pax6b* mutants (with Fold-Discovery-Rate < 10%). The significance in the overlap of the sets of genes regulated by Pax6b in EECs and in PECs was determined with the R built-in statistical Fisher's Exact Test.

Gene Ontology enrichment analysis.

Gene Ontology (GO) enrichment analysis was performed on the different gene sets (genes selectively expressed either in EECs or in PECs, pax6b-regulated genes only in PECs, only in EECs and in both PECs and EECs) using the DAVID bioinformatics resources 6.8 [71] taking as background all the zebrafish genes. The enrichment analysis was focused on the GO biological process, molecular function and KEGG pathways with a statistical Fisher exact test p-value<0.05.

***In situ* hybridization and immunohistochemistry.**

Antisense RNA probe for the different genes were prepared as described by Thisse et al 2008 [72]. Briefly, primers were designed to amplify a part of the transcript that is used as a template to synthesize the probe. The reverse primer at the 5' contains the minimal promoter sequence for T3 RNA polymerase (5'-AATTAACCCTCACTAAAGGGAG-3'), templates were amplified by RT-PCR. Whole mount *in situ* hybridization and fluorescent *in situ* hybridization (WISH and FISH) were performed as described previously [73] on wild-type (AB strain), pax6sa0086 null

mutant or pax6bsunrise mutant embryos/larvae [13]. Immunohistochemistry (IHC) on whole-mount embryos/larvae was performed as described [74]. The antibodies used are the guinea pig polyclonal anti-PDX1 antibody (ab47308, abcam) used at a dilution of 200X, the chicken anti-GFP (Aves lab) used at a dilution of 1000 X, the rabbit anti-Enkephalin (T4294 ; Peninsula Laboratories International, Inc) used at a 400X dilution, the rabbit anti-PACAP (Adcyap1)(T-4473.0050; from Peninsula Laboratories International, Inc) used at a 300X dilution, the monoclonal mouse Hu antibody (16A11 : invitrogen Cat. #A-21271) used at a 1000X dilution, and Alexa Fluor secondary antibodies at 1000X dilution. Stained embryos were mounted in Prolong (Invitrogen) with DAPI and imaged using SP5 confocal microsope (Leica). For the immunostainings on adult tissue, the intestine of transgenic *Tg(-8.5nkx2.2a:GFP)*^{ia2} adult zebrafish was dissected and fixed for 1 day in PFA 4% rinsed in PBS and embedded in TEK using standart procedures and sectioned using cryostat for generating 10µm sections which were mounted on glass slides. The primary and secondary antibodies were used as the same dilution as described above for whole-mount larvae.

List of abbreviations

EEC : enteroendocrine cell

PEC : pancreatic endocrine cell

EN : enteric neuron

WISH : whole-mount *in situ* hybridization

FISH : Fluorescent *in situ* hybridization

CPM : count of reads per million

FPKM : fragments per kilobase of exon model per million reads mapped

S.E : standard errors

Gcg :Glucagon

Sst : Somatostatin

Ghrl : Ghrelin

Pyy : Peptide YY

Ckk : Cholecystokinin

Gip : Glucose-dependent insulinotropic polypeptide

Vip : Vasoactive intestinal peptide

mlnl : Motilin-like

Insl5 : Insulin-like peptide 5

Nmba : Neuromedin B a

Adcyap1 : Adenylate cyclase-activating polypeptide 1

Galn : Galanin

Calca :Calcitonin

Penka : Proenkephalin-a

Declarations

All animal experiments were conducted according to national guidelines and were approved by the ethical committee of the University of Liège (protocol numbers 1328 and 1557). All the analysed data (expression values for all genes in wild-type and mutant EECs and PECs) are given in Tables S1, S4 and S5. The raw datasets have been deposited on Gene Expression Omnibus (GEO) under the accession number GSE149081 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE149081). The authors declare that they have no competing interests.

Acknowledgments

We are grateful for Christian Helker and Didier Stainier for the zebrafish transgenic line TgBAC(pdx1:EGFP)^{bns13} , M. Takamiya and Uhle Strähle for the sunrise mutant line, Francesco Argenton for the Tg(nkx2.2::GFP) and Kenneth Wallace for the *phox2b* probe. We thank the following technical platforms: GIGA-Zebrafish (H Pendeville), GIGA-Cell Imaging and Flow Cytometry (S Ormenese and S Raafat), GIGA-histology (C Humblet) and the GIGA-Genotranscriptomic platforms (W Coppieters, and L Karim).

AL, JP and A-SR were supported by FRIA (Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture) and the Léon Fredericq fund. ETS was supported by WBI (Wallonie-Bruxelles International), Becas Chile and the Léon Fredericq fund. This work was funded by the FNRS-FRS, the ZENCODE-ITN european project 643062, and the Fonds Speciaux from the ULiège (University of Liège). BP, IM, and MLV are associate researchers from FRS/FNRS (Fonds National pour la Recherche Scientifique).

Authors' contributions

ETS and JP carried out the RNA-seq experiments. AL and ETS performed the RNA-seq analyses. AL, ASR and MLV performed the WISH and FISH experiments. AL carried out the bioinformatics analyses concerning the transcriptome comparison, differential gene expression and cell subtype markers. BP performed the IF. BP and MLV conceived the study. BP, AL, MLV and IM participated in the interpretation of the data and wrote the manuscript. All authors read and approved the final manuscript.

AL, ETS, MLV and BP contributed equally to this work.

Figure Legends.

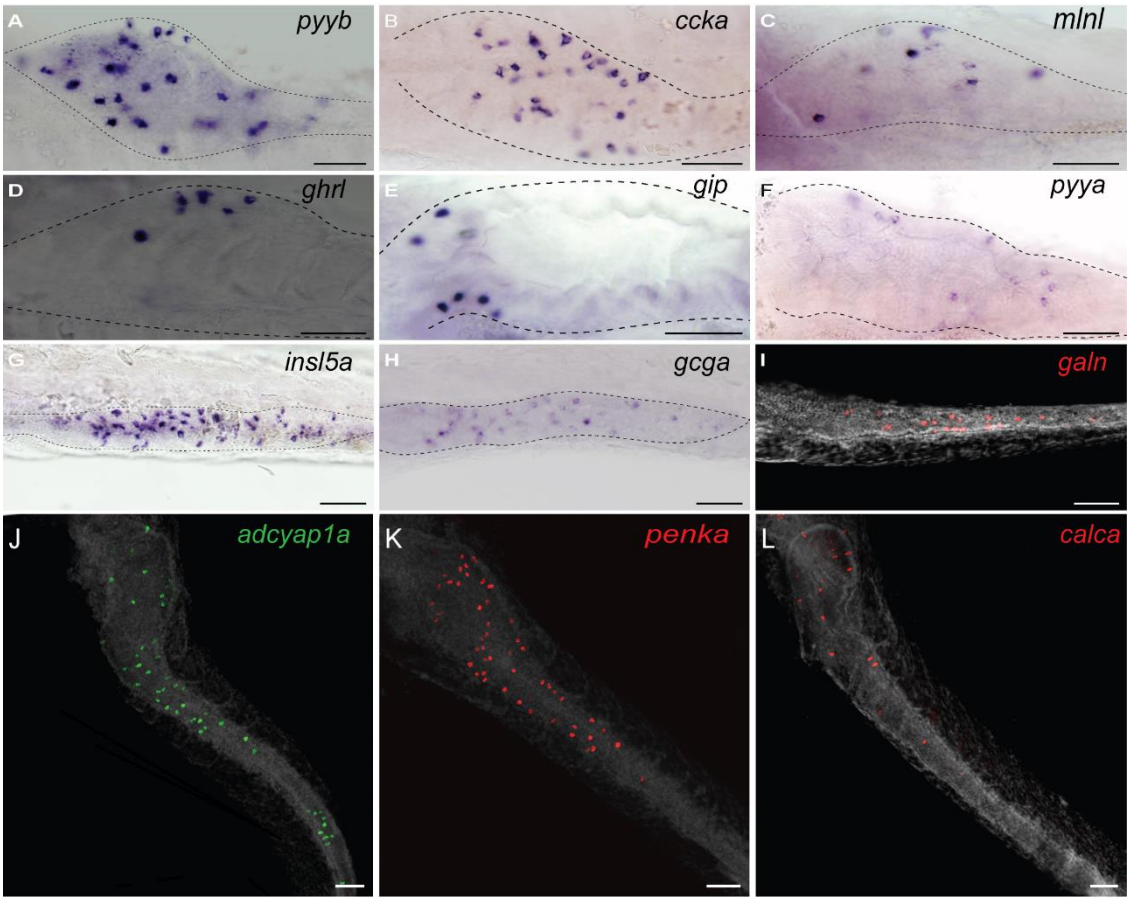


Figure 1: Expression of different enteroendocrine hormones in the zebrafish gut. Whole-mount *in situ* hybridization (WISH) obtained with different hormonal probes at 3 dpf (A-H and K) or 4 dpf (I and K). Ventral views of embryos with anterior on the left. The dotted lines represent the location of the gut. Probes used are : *sst2: Somatostatin* 2. *adcyap1a: adenylate cyclase activating polypeptide 1a*; *vip2: vasoactive intestinal peptide 2*; *ccka: cholecystokinin a*; *gcga: Glucagon a*. *pyya/b: Peptide YYa/b*. *ghrl*: *Ghrelin*. *gip*: *glucose-dependent insulinotropic polypeptide*. *insl5a: Insulin-like peptide 5a*. *calca: Calcitonin*. Scale bars: 50 µm. Panels I to L : fluorescent *in situ* hybridization (FISH)(DAPI staining shown in grey).

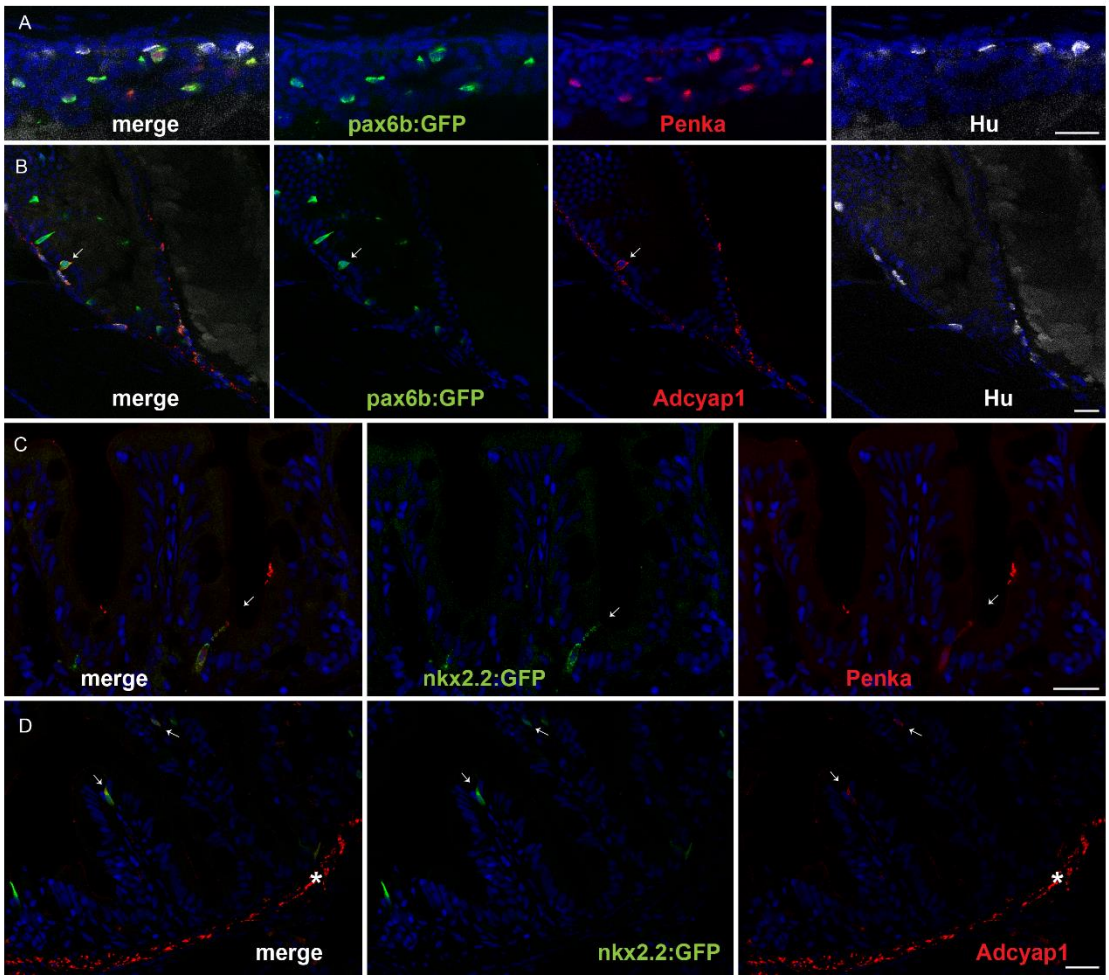


Figure 2: Immunostaining of Enkephalin and Adcyap1 in the zebrafish intestine.

(A-B) Confocal images of intestines from 5-dpf *Tg(pax6b:GFP)* zebrafish larva stained with antibodies directed against GFP (green), Hu (white), ENKEPHALIN (A; red) or ADCYAP1 (B; red). (C-D) Confocal images of intestine of adult *Tg(8.5nkx2.2:GFP)* stained with antibodies directed against GFP(green), ENKEPHALIN (C; red) or ADCYAP1 (D;red). Arrows show location of some EECs expressing Adcyap1 (B and D) or Enkephalin (C). Adcyap1 immunostaining is also observed at level of EN axon fibers in gut larva (B) and in the muscularis layer of adult intestine (shown by * in D). The zebrafish Adcyap1a and Adcyap1b display respectively 6 and 3 mismatches with human ADCYAP1 at the level of the epitope reacting with the Adcyap1 antibody, explaining the stronger signal of Adcyap1b in EN fibers. (The nkx2.2:GFP transgene is expressed specifically in EECs like Pax6b). Scale Bar: 20µm

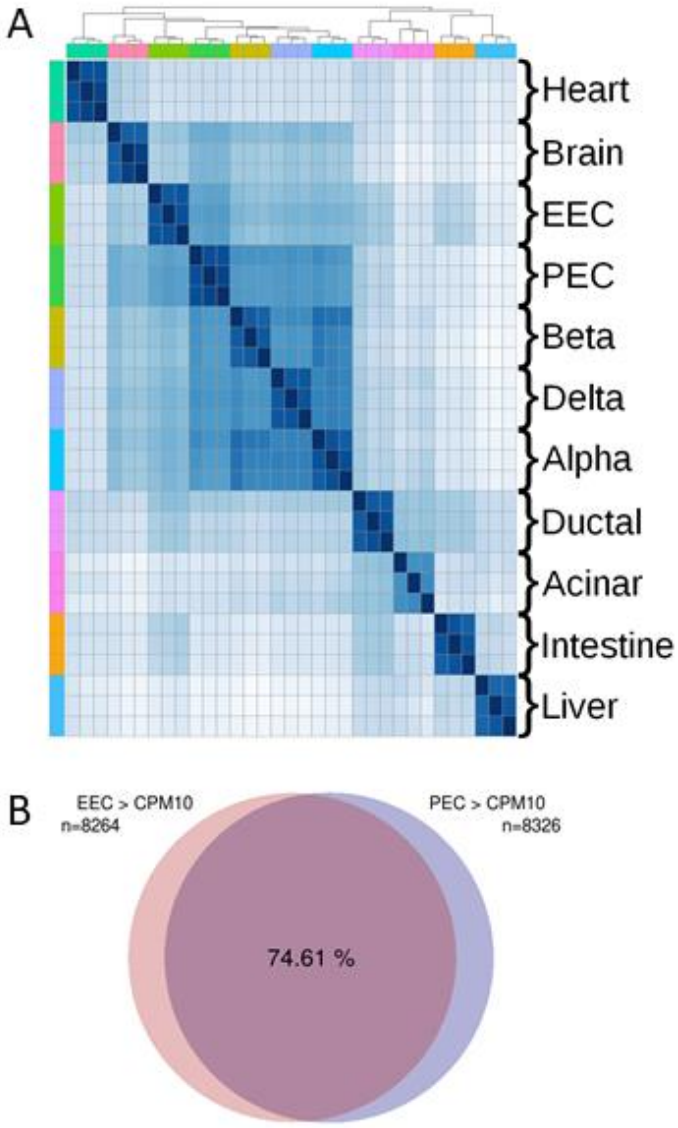


Figure 3: Similar gene expression profiles between zebrafish EECs and PECs.

(A) Global comparison of transcriptomic profiles (RNA-seq) from distinct zebrafish tissues/organs. Clustered heatmap displaying Euclidean distance matrix between every pair of zebrafish RNA-seq datasets. The order of the different tissues is identical for the two axes (rows and columns) as shown by the different colors on the upper part and left part of the matrix. Each zebrafish tissue (analysed in triplicate) was compared with the other samples as presented in the matrix. Darker colour indicates closer distance (i.e. more similar transcriptomes). EEC transcriptome is mostly similar to the transcriptome of embryonic pancreatic endocrine cells (PECs) or to the adult alpha, beta or delta pancreatic cells. The accession number of all used RNA-seq datasets are given in Materials and methods (B) Venn diagram showing the overlap (74%) of genes expressed in both EECs and PECs above the threshold of 10 CPM. 8264 genes and 8326 genes were expressed above 10 CPM (counts per million) respectively in EECs and in PECs.

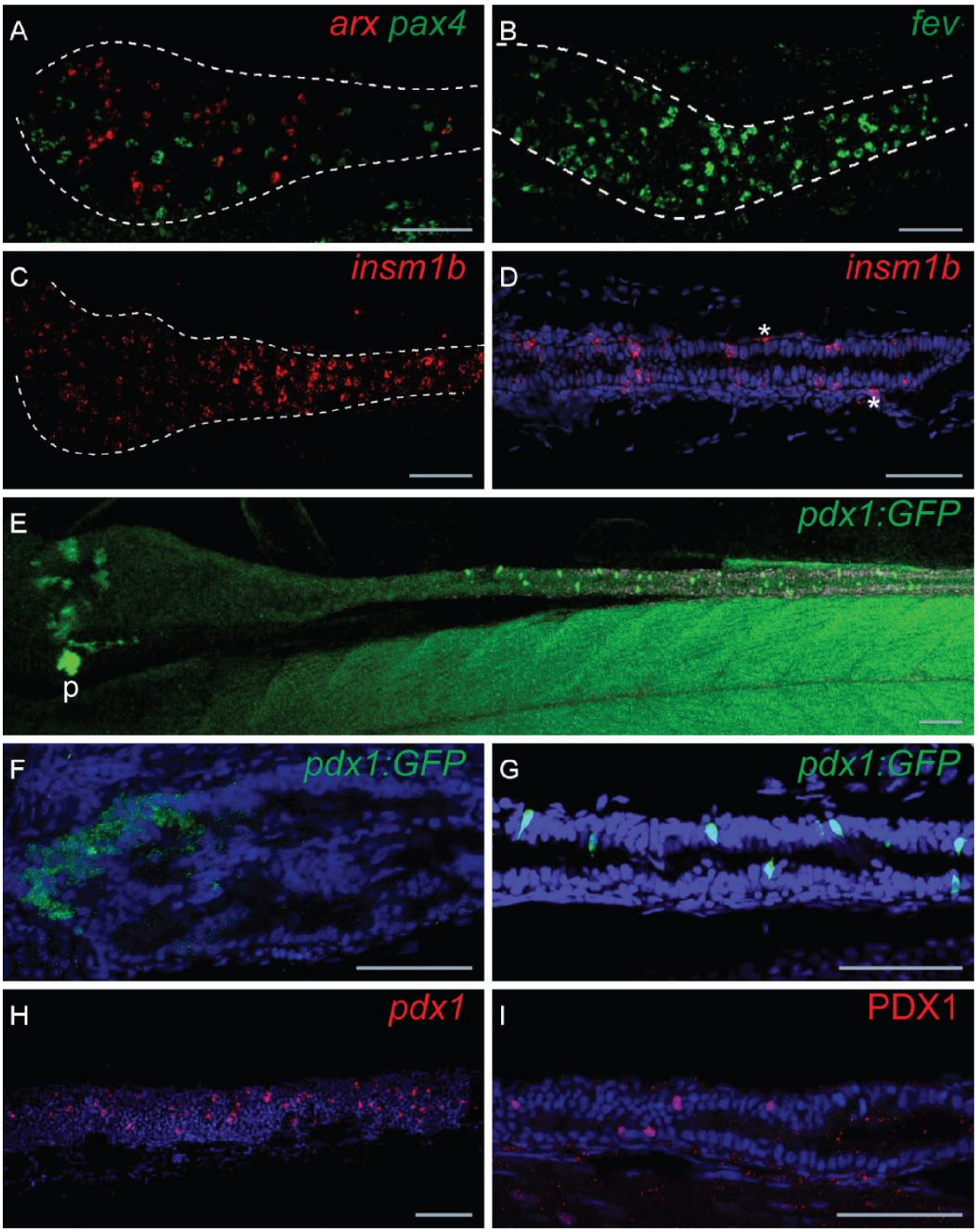


Figure 4: Expression profiles of *pax4*, *arx, pdx1* and *fev* in the zebrafish gut.

Fluorescent in situ hybridization (FISH) performed on 66hpf (A), 72hpf (B-D) or 96 hpf (E-I) embryos. Views of the gut (delimited by dotted lines; anterior to the left) showing the non-overlapping expression of *pax4* (green) and *arxa* (red) (A), numerous EECs expressing *fev* (B) and *insm1b*. (D) Higher magnification showing *insm1*+ EECs located within the gut epithelium and *insm1*+ enteric neurons (shown by asterisks) outside but juxtaposed to the epithelium. (E) View of the gut from the tg(BAC pdx1:GFP) larvae (GFP immunostaining) with higher magnification of the anterior (F) and posterior (G) parts. FISH using *pdx1* probe (H) and immunostaining using Pdx1 antibody. P: pancreas. Scale bars: 20µM.

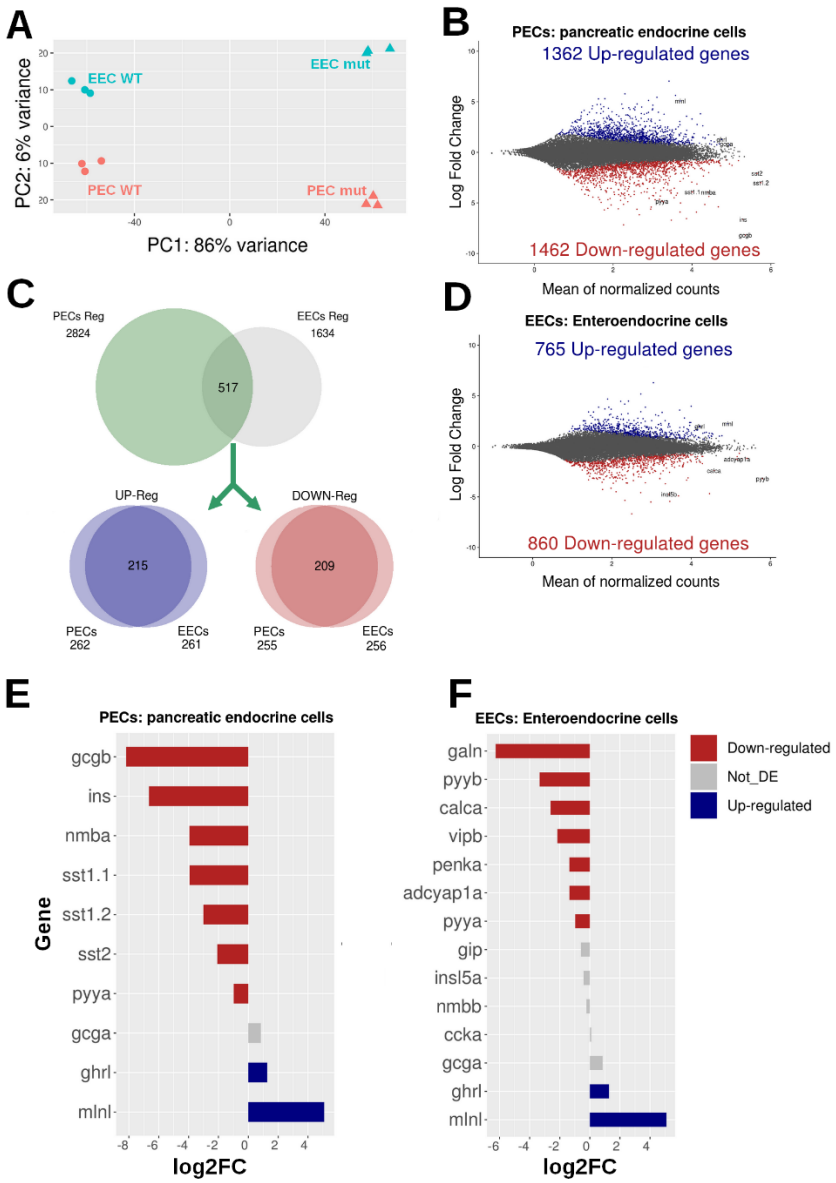


Figure 5: RNA-seq analysis of EECs and PECs from wild-type and *pax6b* mutants identifies a set of genes displaying similar Pax6b-regulation in both pancreas and intestine.

(A) Principal component analysis (PCA) on all EEC and PEC RNA-seq data obtained from wild-type and pax6b^{sa0086} mutants. The close clustering of the triplicate wild-type and mutant samples demonstrates the high reproducibility of the data. (B, D) MA plots showing all the up-regulated (in blue) and down-regulated (in red) genes in the *pax6b*^{sa0086} mutant PECs (B) and EECs (D). (C) Venn diagrams showing the overlap of genes regulated in both EECs and PECs; among the 517 regulated genes, 215 and 209 genes are respectively up- and down- regulated in both cell types as shown in the blue and red Venn diagrams. (E,F) Expression ratio (in log2 of Fold Change from *pax6b* mutant versus wild-type) of hormones expressed in PECs (E) and in EECs (F); down-regulated hormones are in red, up-regulated hormones are in blue, and not statistically affected in grey; RNA-seq values are shown in Tables 3 and S4.

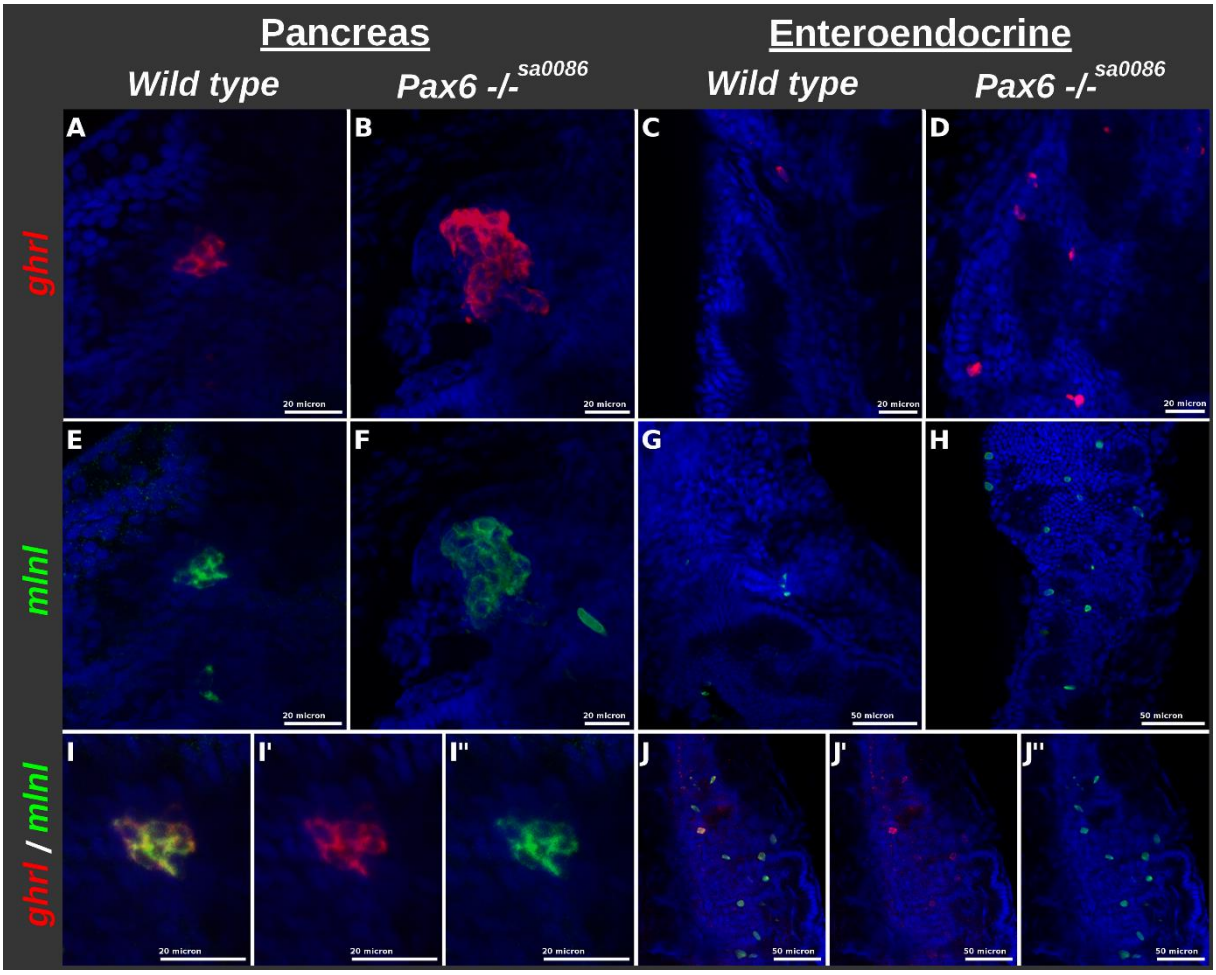


Figure 6: Increase in the number of cells co-expressing *ghrl* and *mlnl* in the *pax6b-/-* PECs and EECs.

FISH using ghrelin (ghrl, labelled in red) and motilin-like (mlnl, labelled in green) probes on 2.5 dpf (A,B,E,F and I) and 4.5 dpf (C,D,G,H,J) zebrafish larvae. Confocal views of the pancreatic islet (A,B,E,F) and of the gut (C,D,G,H). I and J images are overlays of *ghr*l (I' and J') and *mlnl* (I'' and J'') stainings demonstrating co-labelling with the two probes. A,E, I, I' and I'' pictures are views of the pancreatic islet from a wild-type larva while B and F are pancreatic islet from *pax6b-/-*. C and G are views of intestines from wild-type larvae; D, H, J, J' and J'' show views of the gut from a *pax6b-/-* larvae.

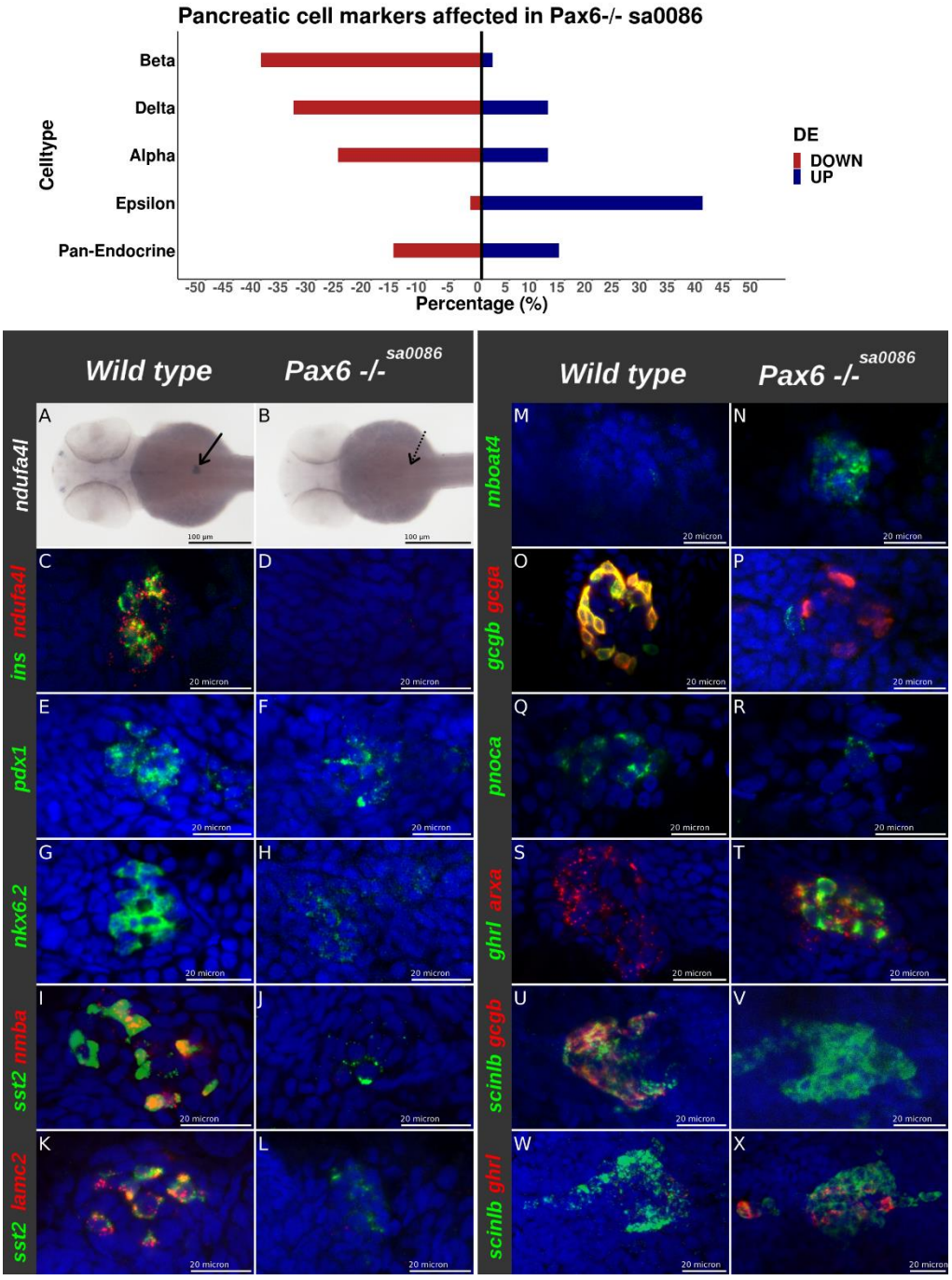


Figure 7: *pax6b* inactivation leads to a loss of pancreatic beta-cells, reduction of delta-cells, increase of epsilon-cells and abnormal alpha-cells.

(Upper panel) percentage of markers from beta-, delta-, alpha-, epsilon- and pan- endocrine cells which are down-regulated (in red) or up-regulated (in blue) in the pax6b-/- PECs. The list of markers for each cell subtype is shown in Table S6. The up- or down- regulation was determined from RNA-seq data from pax6b^{sa0086} mutant versus wild-type. (Lower panel) Expression analysis of cell subtype markers by WISH (A,B) and FISH (C-X) on 2dpf zebrafish *pax6b* mutant and wild-type embryos as noted on the top of the panel; probes are indicated on the left of each picture.

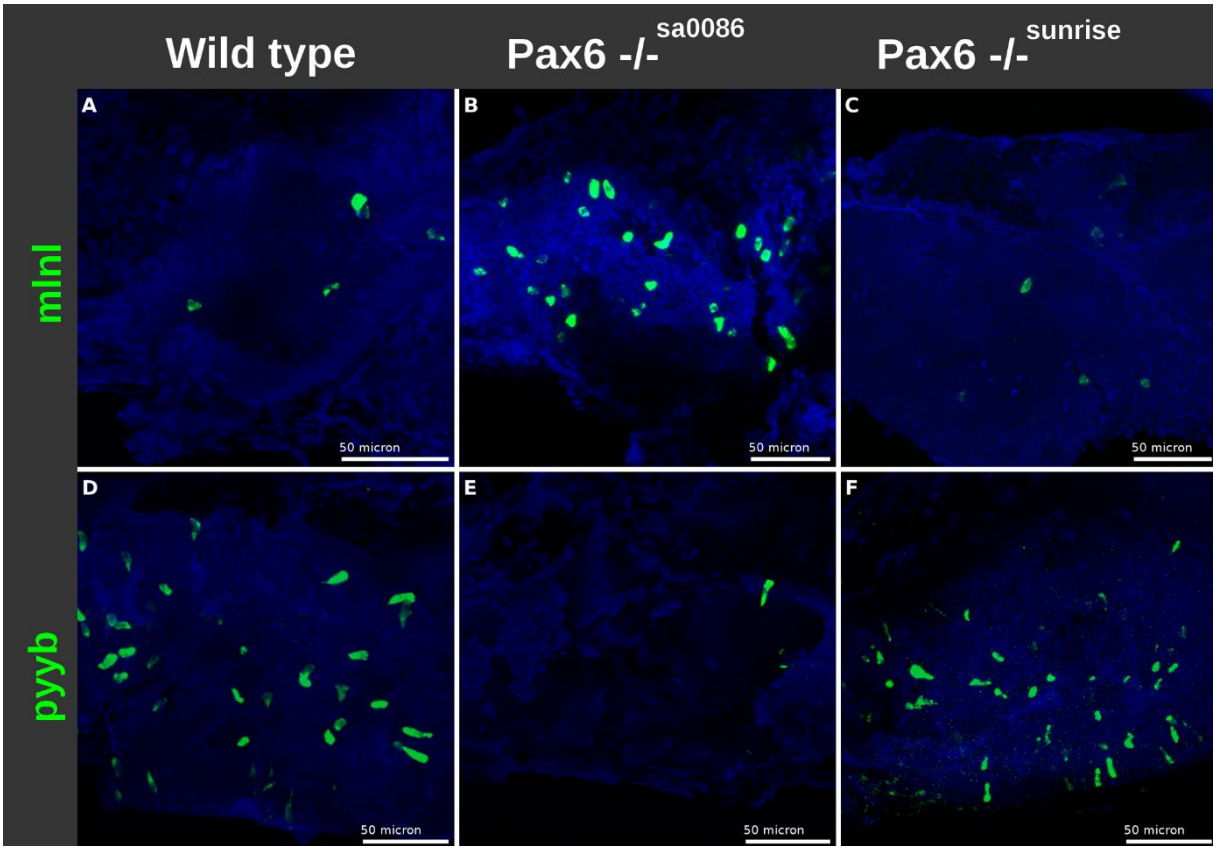


Figure 8: The effect of Pax6b on the number of EEC expressing *pyyb* and *mlnl* does not depend on its homeodomain.

FISH showing the respective increase of *mlnl*+ EECs and decrease of *pyyb*+ EECs in the *pax6b*^{sa0086} null mutants (B,E), while the hypomorphic *pax6b*^{sunrise} mutants harbouring a mutation in *pax6b* homeobox do not display modifications in the expression of EEC hormones (C,F).

Supplemental figure legends

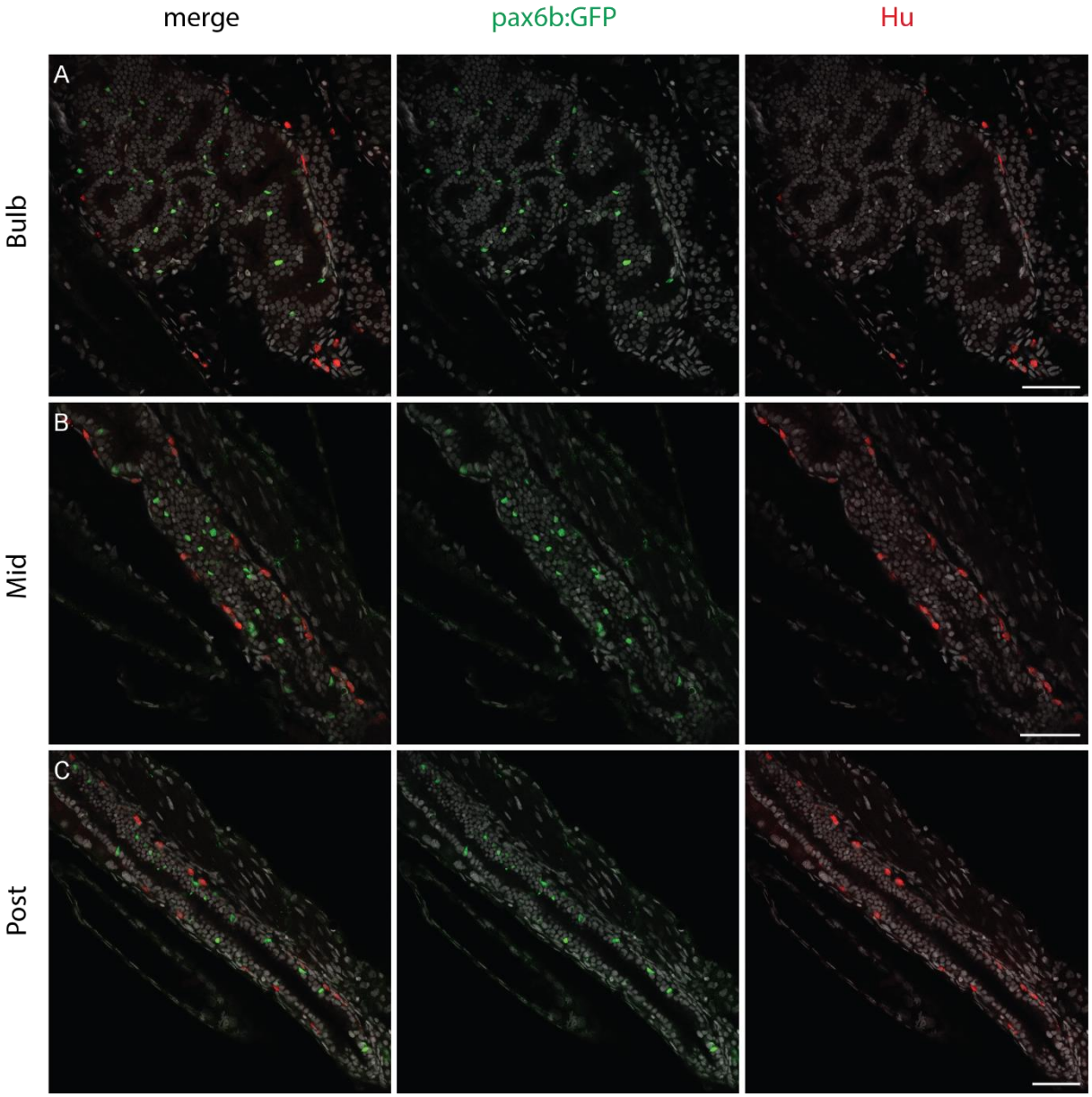


Figure S1: The pax6b:GFP transgene is not expressed in enteric neurones.
Immunofluorescence on 5 dpf *Tg(pax6b:GFP)* larva using antibodies against GFP (green) and against the enteric neurone marker Hu (red). Confocal vues of the bulb intestine (upper panels), mid-intestine (middle panels) and posterior intestine (lower panels) showing no colocalisation of GFP with Hu. Dapi staining is in blue. Scale bar =50µm.

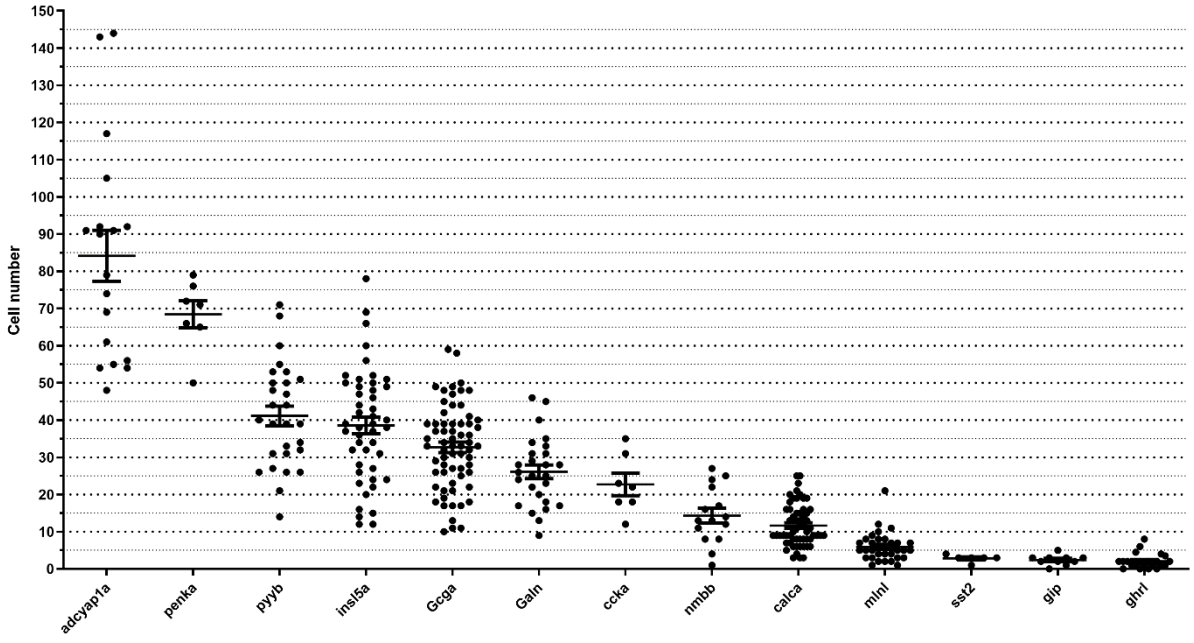


Figure S2: Quantification of EECs expressing different neuropeptide transcripts in 4 dpf zebrafish larvae.

The number of EECs expressing each hormones was determined by counting the labelled cells after WISH using the corresponding hormone probes on 4 dpf larvae. Each point in the graph represents the number of labelled cells in one larva. Bars represent the mean values and S.E.

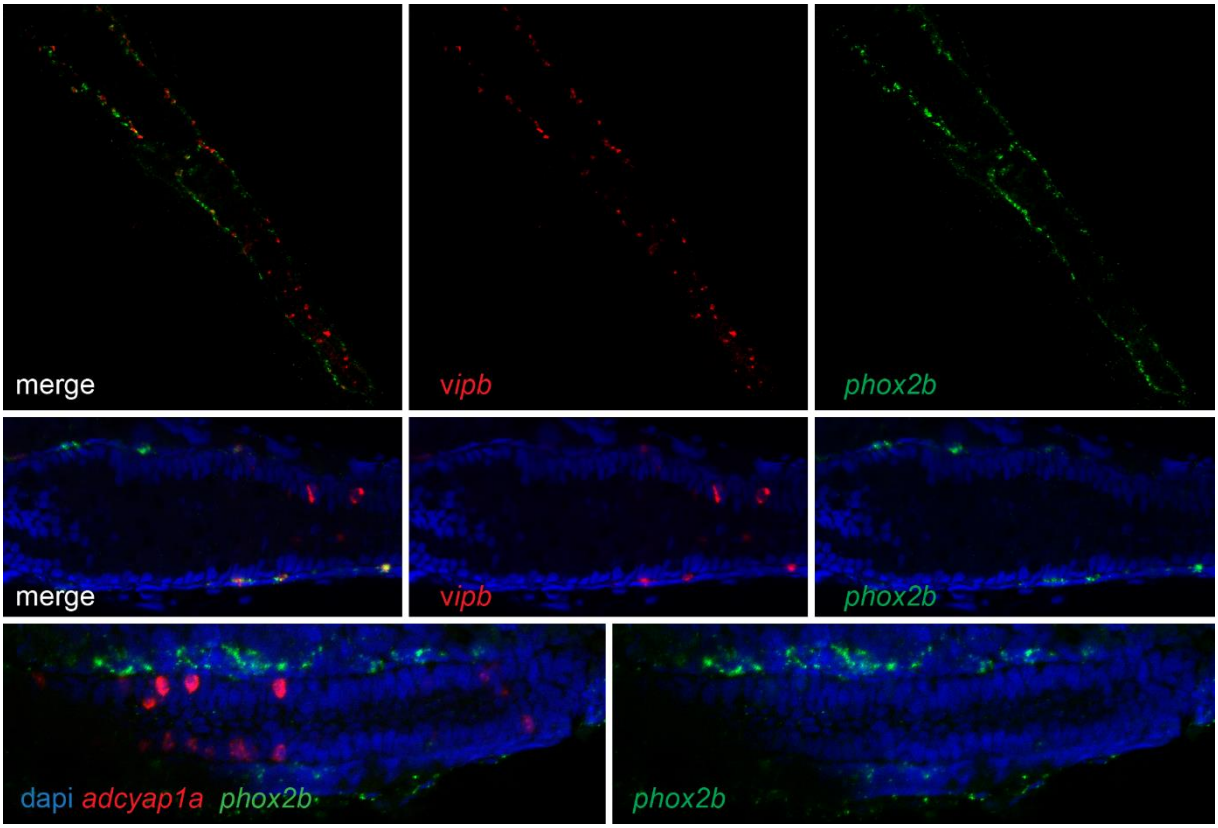


Figure S3: Expression of *vipb* and *adcyap1a* transcripts in zebrafish EECs.

Confocal images of the zebrafish gut from larvae stained by double fluorescent *in situ* hybridization (FISH) using the enteric neurone marker *phox2b* (green) and the *vipb* or *adcyap1a* probes (red). (A) general vue of the gut showing co-localisation of *vipb* and *phox2b* in ENs at the level of anterior intestine (upper left part of the image) and *vipb*+ EECs in the posterior intestine (bottom right of image). (B) higher magnification of the gut showing three vipb+ phox2b+ ENs and two vipb+ EECs. (C) adcyap1a+ cells are distinct from phox2b+ ENs which are located outside the intestinal epithelium (Dapi staining in blue).

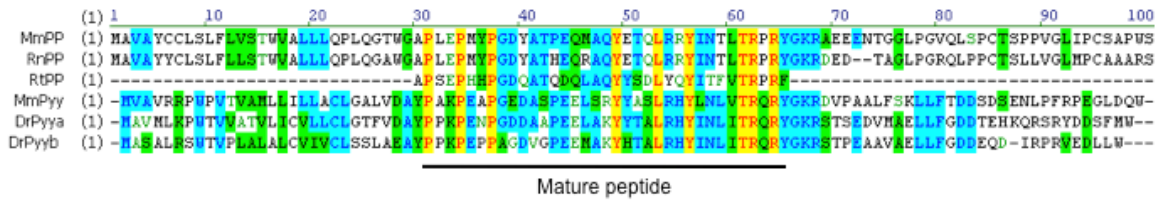


Figure S4 : Sequence similarity between pancreatic polypeptide and of peptide YY.

Alignment of the amino-acid sequence of pancreatic polypeptide (PP) from mice and rat with peptide YY from mice and zebrafish (DrPyy A or b).

Tables

Table 1: Identification and expression level of zebrafish enteroendocrine hormones

	Mammalian Hormones		Zebrafish Hormones	Synteny	Onset time	Expression (WISH)			Cell Number mean±SE	EEC RNAseq (CPM)	EN RNAseq (CPM)
						Bulb	Mid gut	Posterior gut			
A	*1*	*Gcg*	*gcga*	yes	3.5 dpf	-	✓	✓	33±1.4	2058	57
			gcgb	yes	N.T					177	14
	2	*Sst*	*sst1.1*	yes	-	-	-	-		20	13
			sst1.2	yes	N.T					88	1519
			sst2	yes	3 dpf	✓	✓	-	2.8±0.4	103	933
	3	*Ghrl*	*ghrl*	no	3 dpf	✓	-	-	2.1±0.4	163	40
	4	*Pyy*	*pyya*	yes	4 dpf	✓	-	-		437	1
			pyyb	yes	3 dpf	✓	✓	-	41±2.6	32011	4
	5	*Cck*	*ccka*	yes	3 dpf	✓	-	-	23±3.0	6390	1
	6	*Gip*	*gip*	yes	4 dpf	✓	-	-	2.4±0.4	100	4
	7	*Vip*	*vip*	yes	N.T					0	511
			vipb	no	4 dpf	-	-	✓	N.D*	927	868
	8	*Nts*	*nts*	yes	-	-	-	-		0	3
	9	*Insl5*	*insl5a*	no	4 dpf	-	✓	✓	39±2.2	5602	14
			insl5b	no	N.T					178	1
	10	*Nmb*	*nmba*	yes	N.T					45	116
			nmbb	no	3 dpf	✓	-	-	14±2.0	2150	0
	11	*Galn*	*galn*	yes	3.5 dpf	-	✓	✓	26±1.8	824	157
	12	*Mln*	*/*	/	/	/	/	/		/	
			mlnl	no	3 dpf	✓	/	/	5.9±0.7	708	18
	13	*Gast*	*/*	/	/	/	/	/		/	
	14	*Sct*	*/*	/	/	/	/	/		/	
	15	*Pp*	*/*	/	/	/	/	/		/	
B	*16*	*Adcyap1*	*adcyap1a*	no	3 dpf	✓	✓	✓	84±6.9	6927	2
			adcyap1b	no	N.T					6	251
	17	*Calca*	*calca*	yes	4 dpf	✓	✓	-	12±0.7	2161	1
	18	*Penk*	*penka*	yes	3 dpf	✓	✓	-	68±3.6	12544	2
			penkb	yes	N.T					2	

(Upper panel, lanes 1-15) Identification of the zebrafish orthologs for the 15 mammalian genes coding for enteroendocrine hormones by “in silico” search of the zebrafish genome. (Lower panel, lanes 16-18) Identification of the 3 novel hormones expressed in zebrafish EECs by screening the RNA-seq data with the Gene Ontology term “hormone activity”(GO: 0005179) and selecting hormones expressed at high levels (>2.000 CPM). Column 3 indicates the Ensembl ID of the zebrafish genes. Column 4 displays whether a synteny conservation is observed between mouse and zebrafish loci. Column 5 indicates the expression onset time determined by WISH for each hormone transcripts and their localization in the zebrafish gastrointestinal tract (column 6). Bulb corresponds to the enlarged rostral part of the intestine; the mid-gut is located posterior to the bulb and contains the majority of goblet cells at 5 dpf; the posterior gut corresponds the most caudal part of intestine containing few goblet cells [29]. Column 7 indicates the mean (± standard errors (SE) of the EEC numbers detected by WISH at 4 dpf. The number of EEC counted for each individual larvae is provided in Fig. S2. N.D* : not determined; the number of *vipb*+ EEC cells could not be quantified as most of the *vipb*+ cells detected in the gut are *vipb*+ enteric neurons. Column 8 indicates the expression levels determined by RNAseq of EEC in CPM (counts per million). Column 9 indicates the expression levels determined by RNAseq of enteric neurons in CPM [33]. – : not detected by WISH; v : detected by WISH.

The abbreviations used are Glucagon (Gcg), Somatostatin (Sst), Ghrelin (Ghrl), Peptide YY (Pyy), Gastrin (Gast), Cholecystokinin (Ckk), Glucose-dependent insulinotropic polypeptide (Gip), Vasoactive intestinal peptide (Vip), Motilin (Mln), Motilin-like (mlnl), Pancreatic Polypeptide (Pp), Neurotensin (Nts), Secretin (Sct), Insulin-like peptide 5 (Insl5), Neuromedin B (Nmb), Adenylate cyclase-activating polypeptide 1 (Adcyap1), Galanin (Galn), Calcitonin (Calca).

Table 2: Expression level in EECs of transcription factors involved in zebrafish PEC differentiation.

Pancreatic Endocrine TF	Onset time in the gut	Expression in the gut			EEC RNAseq (CPM)	EN RNAseq (CPM)
		Bulb	Mid gut	Posterior gut		
ascl1b	-	-	-	-	2	3
*ascl1a**	36hpf	✓	✓	✓	25	55
*sox4b**	38hpf	✓	✓	✓	36	22
insm1b	≤52hpf**	✓	✓	✓	638	189
insm1a	≤52hpf**	N.D	N.D	N.D	123	38
foxo1a	≤52hpf**	✓	✓	✓	504	33
*neurod1**	52hpf	✓	✓	✓	474	67
*nkx2.2a**	52hpf	✓	✓	✓	195	2
pax4	58hpf	✓	✓	✓	225	0
*pax6b**	60hpf	✓	✓	✓	826	23
*rfx6**	62hpf	✓	-	-	13	8
fev	62hpf	✓	✓	✓	951	3
*isl1**	65hpf	✓	✓	✓	134	31
arxa	66hpf	✓	✓	✓	54	1
pdx1	72hpf in EEC	-	✓	✓	614	4
nkx6.1	-	-	-	-	1	2
nkx6.2	-	-	-	-	5	1
mnx1	-	-	-	-	1	0

Recapitulation of the WISH data performed on zebrafish embryos from 36 to 72hpf showing the expression onset time and the distribution profile in the gut determined by WISH for 17 transcription factors known to be involved in pancreatic endocrine cell (PEC) differentiation. **Based on RNAseq data from FACS-sorted sox17:dsRED cells of the microdissected intestine at 52 hpf, *insm1a, insm1b* and *foxo1a* are already expressed in the gut at 52 hpf (Reuter *et al.,* article in preparation). The two last columns indicate the expression level in EECs and in enteric neurons (ENs) as determined by RNA-seq in normalized CPM. (N.D.: not determined; - : not detected by WISH; v : detected by WISH). All these TF are expressed at significant levels in EEC, except *ascl1b*, *nkx6.1*, *nkx6.2* and *mnx1*. Expression data of genes labelled with * are from [27].

Table 3: Expression level of some selected markers in EEC and PEC isolated from wild-type and pax6b^{sa0086} mutant larvae.

markers	Gene	PEC wildtype	PEC pax6b-/-	EEC wildtype	EEC pax6b-/-
epsilon PEC	***ghrl***	996	3156	184	876
	mlnl	7,6	385	785	4604
	mboat4	0,7	201	0,02	1,3
beta PEC	***ins***	13111	112	1,1	2,9
	ndufa4l2a	1729	82	1,8	1,2
	nkx6.2	330	10	5,5	7,7
	pdx1	1377	37	453	46
delta PEC	***sst2***	24070	6287	107	440
	sst1.1	668	37	27	0,2
	sst1.2	33282	4375	89	37
	lamc2	153	16	4	31
	nmba	1640	110	46	15
	hhex	95	27	0,5	0,1
	cdx4	181	64	40	32
alpha PEC	***gcgb***	15002	42	179	2,4
	gcga	1776	3791	1982	1526
	pnoca	144	14	11	2,2
alpha and epsilon	***scinlb***	83	1374	476	1365
alpha and epsilon	***arxa***	67	290	42	57
EEC hormones	***galn***	0,05	0,19	758	12
	pyyb	32	20	34323	3746
	calca	57	24	1933	349
	adcyap1a	0,58	0,41	6201	2672
	vipb	0,07	0,17	737	190
	ccka	0,43	1,01	6257	7212

The expression level (given in normalized CPM) was obtained from the RNA-seq data. The values are the expression mean of triplicate samples. Table S4 provides the expression level of all genes (values for each samples, means and standard deviation).

References

1. Gribble FM, Reimann F. Enteroendocrine Cells: Chemosensors in the Intestinal Epithelium. Annu Rev Physiol. 2016;78:277–99. doi:10.1146/annurev-physiol-021115-105439.

2. Habib AM, Richards P, Cairns LS, Rogers GJ, Bannon CAM, Parker HE, et al. Overlap of endocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. Endocrinology. 2012;153:3054–65. doi:10.1210/en.2011-2170.

3. Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, et al. A single-cell survey of the small intestinal epithelium. Nature. 2017;551:333–9. doi:10.1038/nature24489.

4. Glass LL, Calero-Nieto FJ, Jawaid W, Larraufie P, Kay RG, Göttgens B, et al. Single-cell RNA-sequencing reveals a distinct population of proglucagon-expressing cells specific to the mouse upper small intestine. Mol Metab. 2017;6:1296–303. doi:10.1016/j.molmet.2017.07.014.

5. Böttcher G, Sjöberg J, Ekman R, Håkanson R, Sundler F. Peptide YY in the mammalian pancreas: immunocytochemical localization and immunochemical characterization. Regul Pept. 1993;43:115–30. doi:10.1016/0167-0115(93)90146-Y.

6. Wierup N, Svensson H, Mulder H, Sundler F. The ghrelin cell: a novel developmentally regulated islet cell in the human pancreas. Regul Pept. 2002;107:63–9. doi:10.1016/S0167-0115(02)00067-8.

7. Naya FJ, Huang HP, Qiu Y, Mutoh H, DeMayo FJ, Leiter AB, et al. Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. Genes Dev. 1997;11:2323–34. doi:10.1101/gad.11.18.2323.

8. May CL, Kaestner KH. Gut endocrine cell development. Mol Cell Endocrinol. 2010;323:70–5.

9. Gierl MS, Karoulias N, Wende H, Strehle M, Birchmeier C. The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. Genes Dev. 2006;20:2465–78. doi:10.1101/gad.381806.

10. Beucher A, Gjernes E, Collin C, Courtney M, Meunier A, Collombat P, et al. The homeodomain-containing transcription factors Arx and Pax4 control enteroendocrine subtype specification in mice. PLoS One. 2012;7:e36449. doi:10.1371/journal.pone.0036449.

11. Sander M, Neubuser A, Kalamaras J, Ee HC, Martin GR, German MS. Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes Dev. 1997;11:1662–73. http://www.ncbi.nlm.nih.gov/pubmed/9224716.

12. Heller RS, Jenny M, Collombat P, Mansouri A, Tomasetto C, Madsen OD, et al. Genetic determinants of pancreatic epsilon-cell development. Dev Biol. 2005;286:217–24. doi:10.1016/j.ydbio.2005.06.041.

13. Verbruggen V, Ek O, Georlette D, Delporte F, Von Berg V, Detry N, et al. The Pax6b homeodomain is dispensable for pancreatic endocrine cell differentiation in zebrafish. J Biol Chem. 2010;285:13863–73. doi:10.1074/jbc.M110.108019.

14. Larsson LI, St-Onge L, Hougaard DM, Sosa-Pineda B, Gruss P. Pax 4 and 6 regulate gastrointestinal endocrine cell development. Mech Dev. 1998;79:153–9. http://www.ncbi.nlm.nih.gov/pubmed/10349628. Accessed 29 Jun 2017.

15. Hill ME, Asa SL, Drucker DJ. Essential requirement for Pax6 in control of enteroendocrine proglucagon gene transcription. Mol Endocrinol. 1999;13:1474–86. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10478839.

16. Hartenstein V, Takashima S, Hartenstein P, Asanad S, Asanad K. bHLH proneural genes as cell fate determinants of entero-endocrine cells, an evolutionarily conserved lineage sharing a common root with sensory neurons. Dev Biol. 2017;431:36–47. doi:10.1016/j.ydbio.2017.07.013.

17. Hartenstein V, Martinez P. Structure, development and evolution of the digestive system. Cell Tissue Res. 2019;377:289–92. doi:10.1007/s00441-019-03102-x.

18. Annunziata R, Andrikou C, Perillo M, Cuomo C, Arnone MI. Development and evolution of gut structures: from molecules to function. Cell Tissue Res. 2019;377:445–58. doi:10.1007/s00441-019-03093-9.

19. Heller RS. The Islets of Langerhans. 2010;654:21–37. doi:10.1007/978-90-481-3271-3.

20. Youson JH, Al-Mahrouki AA. Ontogenetic and Phylogenetic Development of the Endocrine Pancreas (Islet Organ) in Fishes. Gen Comp Endocrinol. 1999;116:303–35. doi:10.1006/gcen.1999.7376.

21. Arendt D. The evolution of cell types in animals: emerging principles from molecular studies. Nat Rev Genet. 2008;9:868–82. doi:10.1038/nrg2416.

22. Kin K, Nnamani MC, Lynch VJ, Michaelides E, Wagner GP. Cell-type phylogenetics and the origin of endometrial stromal cells. Cell Rep. 2015;10:1398–409. doi:10.1016/j.celrep.2015.01.062.

23. Arntfield ME, van der Kooy D. β-Cell evolution: How the pancreas borrowed from the brain: The shared toolbox of genes expressed by neural and pancreatic endocrine cells may reflect their evolutionary relationship. BioEssays. 2011;33:582–7.

24. Tarifeño-Saldivia E, Lavergne A, Bernard A, Padamata K, Bergemann D, Voz MLML, et al. Transcriptome analysis of pancreatic cells across distant species highlights novel important regulator genes. BMC Biol. 2017;15:21.

doi:10.1186/s12915-017-0362-x.

25. Bates JM, Mittge E, Kuhlman J, Baden KN, Cheesman SE, Guillemin K. Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. 2006;297:374–86. https://www.sciencedirect.com/science/article/pii/S0012160606007743?via%3Dihub. Accessed 21 Aug 2019.

26. Chen YH, Lu YF, Ko TY, Tsai MY, Lin CY, Lin CC, et al. Zebrafish cdx1b regulates differentiation of various intestinal cell lineages. Dev Dyn. 2009;238:1021–32.

27. Flasse LCLC, Stern DGDG, Pirson JLJL, Manfroid I, Peers B, Voz MLMLML. The bHLH transcription factor Ascl1a is essential for the specification of the intestinal secretory cells and mediates Notch signaling in the zebrafish intestine. Dev Biol. 2013;376:187–97. doi:10.1016/j.ydbio.2013.01.011.

28. Roach G, Heath Wallace R, Cameron A, Emrah Ozel R, Hongay CF, Baral R, et al. Loss of ascl1a prevents secretory cell differentiation within the zebrafish intestinal epithelium resulting in a loss of distal intestinal motility. Dev Biol. 2013;376:171–86. doi:10.1016/j.ydbio.2013.01.013.

29. Wallace KN, Akhter S, Smith EM, Lorent K, Pack M. Intestinal growth and differentiation in zebrafish. Mech Dev. 2005;122:157–73.

30. May CL, Kaestner KH. Gut endocrine cell development. Mol Cell Endocrinol. 2010;323:70–5. doi:10.1016/j.mce.2009.12.009.

31. Liu Y, Li S, Huang X, Lu D, Liu X, Ko WH, et al. Identification and characterization of a motilin-like peptide and its receptor in teleost. Gen Comp Endocrinol. 2013;186:85–93.

32. Delporte FM, Pasque V, Devos N, Manfroid I, Voz ML, Motte P, et al. Expression of zebrafish pax6b in pancreas is regulated by two enhancers containing highly conserved cis-elements bound by PDX1, PBX and PREP factors. BMC Dev Biol. 2008;8.

33. Roy-Carson S, Natukunda K, Chou H, Pal N, Farris C, Schneider SQ, et al. Defining the transcriptomic landscape of the developing enteric nervous system and its cellular environment. BMC Genomics. 2017;18:290. doi:10.1186/S12864-017-3653-2.

34. Uyttebroek L, Shepherd IT, Harrisson F, Hubens G, Blust R, Timmermans JP, et al. Neurochemical coding of enteric neurons in adult and embryonic zebrafish (Danio rerio). J Comp Neurol. 2010;518:4419–38.

35. van Arensbergen J, García-Hurtado J, Moran I, Maestro MA, Xu X, Van de Casteele M, et al. Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. Genome Res. 2010;20:722–32. doi:10.1101/gr.101709.109.

36. Collombat P, Mansouri A, Hecksher-Sorensen J, Serup P, Krull J, Gradwohl G, et

al. Opposing actions of Arx and Pax4 in endocrine pancreas development. Genes Dev. 2003;17:2591–603. doi:10.1101/gad.269003.

37. Djiotsa J, Verbruggen V, Giacomotto J, Ishibashi M, Manning E, Rinkwitz S, et al. Pax4 is not essential for beta-cell differentiation in zebrafish embryos but modulates alpha-cell generation by repressing arx gene expression. BMC Dev Biol. 2012;12:37. doi:10.1186/1471-213X-12-37.

38. Ohta Y, Kosaka Y, Kishimoto N, Wang J, Smith SB, Honig G, et al. Convergence of the insulin and serotonin programs in the pancreatic β-cell. Diabetes. 2011;60:3208–16. doi:10.2337/db10-1192.

39. Helker CSM, Mullapudi S-T, Mueller LM, Preussner J, Tunaru S, Skog O, et al. A whole organism small molecule screen identifies novel regulators of pancreatic endocrine development. Development. 2019;146:dev172569. doi:10.1242/dev.172569.

40. Yee NS, Lorent K, Pack M. Exocrine pancreas development in zebrafish. Dev Biol. 2005;284:84–101. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15963491.

41. Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, Ninov N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. Nat Biotechnol. 2018;36:469–73. doi:10.1038/nbt.4124.

42. Kleinjan DA, Bancewicz RM, Gautier P, Dahm R, Schonthaler HB, Damante G, et al. Subfunctionalization of Duplicated Zebrafish pax6 Genes by cis-Regulatory Divergence. PLoS Genet. 2008;4:e29. doi:10.1371/journal.pgen.0040029.

43. Cardoso JC, Vieira FA, Gomes AS, Power DM. The serendipitous origin of chordate secretin peptide family members. BMC Evol Biol. 2010;10:135. doi:10.1186/1471-2148-10-135.

44. Hort Y, Baker E, Sutherland GR, Shine J, Herzog H. Gene duplication of the human peptide YY gene (PYY) generated the pancreatic polypeptide gene (PPY) on chromosome 17q21.1. Genomics. 1995;26:77–83. doi:10.1016/0888-7543(95)80085-z.

45. Sundström G, Larsson TA, Brenner S, Venkatesh B, Larhammar D. Evolution of the neuropeptide Y family: New genes by chromosome duplications in early vertebrates and in teleost fishes. Gen Comp Endocrinol. 2008;155:705–16. doi:10.1016/J.YGCEN.2007.08.016.

46. Biemar F, Argenton F, Schmidtke R, Epperlein S, Peers B, Driever W. Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. Dev Biol. 2001;230:189–203. doi:10.1006/dbio.2000.0103.

47. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al.

Tissue-based map of the human proteome. Science (80-). 2015;347:1260419–1260419. doi:10.1126/science.1260419.

48. Mutch DM, Anderle P, Fiaux M, Mansourian R, Vidal K, Wahli W, et al. Regional variations in ABC transporter expression along the mouse intestinal tract. Physiol Genomics. 2004.

49. Segerstolpe Å, Palasantza A, Eliasson P, Andersson E-M, Andréasson A-C, Sun X, et al. Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. Cell Metab. 2016;24:593–607. doi:10.1016/j.cmet.2016.08.020.

50. Crosnier C, Vargesson N, Gschmeissner S, Ariza-McNaughton L, Morrison A, Lewis J. Delta-Notch signalling controls commitment to a secretory fate in the zebrafish intestine. Development. 2005;132:1093–104. doi:10.1242/dev.01644.

51. Jensen J, Pedersen EE, Galante P, Hald J, Heller RS, Ishibashi M, et al. Control of endodermal endocrine development by Hes-1. Nat Genet. 2000;24:36–44. doi:10.1038/71657.

52. Hartenstein V, Takashima S, Hartenstein P, Asanad S, Asanad K. bHLH proneural genes as cell fate determinants of entero-endocrine cells, an evolutionarily conserved lineage sharing a common root with sensory neurons. Dev Biol. 2017;431:36–47. doi:10.1016/J.YDBIO.2017.07.013.

53. Falkmer S, Dafgård E, el-Salhy M, Engström W, Grimelius L, Zetterberg A. Phylogenetical aspects on islet hormone families: a minireview with particular reference to insulin as a growth factor and to the phylogeny of PYY and NPY immunoreactive cells and nerves in the endocrine and exocrine pancreas. Peptides. 1985;6 Suppl 3:315–20. doi:10.1016/0196-9781(85)90391-2.

54. Youson JH. The agnathan enteropancreatic endocrine system: Phylogenetic and ontogenetic histories, structure, and function1. Am Zool. 2000;40:179–99.

55. Perillo M, Arnone MI. Characterization of insulin-like peptides (ILPs) in the sea urchin Strongylocentrotus purpuratus: Insights on the evolution of the insulin family. Gen Comp Endocrinol. 2014;205:68–79. doi:10.1016/J.YGCEN.2014.06.014.

56. Ariyachet C, Tovaglieri A, Xiang G, Lu J, Shah MS, Richmond CA, et al. Reprogrammed Stomach Tissue as a Renewable Source of Functional β Cells for Blood Glucose Regulation. Cell Stem Cell. 2016;18:410–21. doi:10.1016/j.stem.2016.01.003.

57. Chen Y-J, Finkbeiner SR, Weinblatt D, Emmett MJ, Tameire F, Yousefi M, et al. De novo formation of insulin-producing "neo-β cell islets" from intestinal crypts. Cell Rep. 2014;6:1046–58. doi:10.1016/j.celrep.2014.02.013.

58. Talchai C, Xuan S, Kitamura T, DePinho RA, Accili D. Generation of functional insulin-producing cells in the gut by Foxo1 ablation. Nat Genet. 2012;44:406–12. doi:10.1038/ng.2215.

59. Thurner M, Bunt M van de, Torres JM, Mahajan A, Nylander V, Bennett AJ, et al. Integration of human pancreatic islet genomic data refines regulatory mechanisms at Type 2 Diabetes susceptibility loci. Elife. 2018;7. doi:10.7554/ELIFE.31977.

60. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet. 2018;50:1505–13. doi:10.1038/s41588-018-0241-6.

61. Scott Heller R, Stoffers DA, Liu A, Schedl A, Crenshaw EB, Madsen OD, et al. The role of Brn4/Pou3f4 and Pax6 in forming the pancreatic glucagon cell identity. Dev Biol. 2004;268:123–34. doi:10.1016/j.ydbio.2003.12.008.

62. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. Dev Dyn. 1995;203:253–310. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8589427.

63. Delporte FMFM, Pasque V, Devos N, Manfroid I, Voz MLML, Motte P, et al. Expression of zebrafish pax6b in pancreas is regulated by two enhancers containing highly conserved cis-elements bound by PDX1, PBX and PREP factors. BMC Dev Biol. 2008;8:53. doi:10.1186/1471-213X-8-53.

64. Pauls S, Zecchin E, Tiso N, Bortolussi M, Argenton F. Function and regulation of zebrafish nkx2.2a during development of pancreatic islet and ducts. Dev Biol. 2007;304:875–90. doi:10.1016/j.ydbio.2007.01.024.

65. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014;9:171–81. doi:10.1038/nprot.2014.006.

66. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21. doi:10.1093/bioinformatics/bts635.

67. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2014;:btu638.

68. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. doi:10.1186/s13059-014-0550-8.

69. Ghaye AP, Bergemann D, Tarifeño-Saldivia E, Flasse LC, Von Berg V, Peers B, et al. Progenitor potential of nkx6.1-expressing cells throughout zebrafish life and during beta cell regeneration. BMC Biol. 2015;13:70. doi:10.1186/s12915-015-0179-4.

70. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.

71. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of

large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57. doi:10.1038/nprot.2008.211.

72. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc. 2008;3:59–69.

73. Mavropoulos A, Devos N, Biemar F, Zecchin E, Argenton F, Edlund H, et al. sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. Dev Biol. 2005;285:211–23. doi:10.1016/j.ydbio.2005.06.024.

74. Flasse LC, Pirson JL, Stern DG, Von Berg V, Manfroid I, Peers B, et al. Ascl1b and Neurod1, instead of Neurog3, control pancreatic endocrine cell fate in zebrafish. BMC Biol. 2013;11:78. doi:10.1186/1741-7007-11-78.

75. Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, Ninov N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nat Biotechnol. 2018;36:469–73. doi:10.1038/nbt.4124.

2.1 Characterization of genomic DNA-binding sites of Pax6b in zebrafish embryos

In order to identify genomic regions bound by Pax6b protein in zebrafish, we performed ChIP-Seq experiments on whole embryos at 27 hpf for two biological replicates. We identified 15792 and 14412 peaks for replicate 1 and 2 respectively, corresponding to genomic region enriched in Pax6b ChIP-seq data. 10811 peaks were detected in the two replicates (Figure 38-A) . The profile of the genomic locations of these peaks is highly similar between replicates with most of them located in intergenic regions (~38%), introns (~46%) and promoter-TSS regions (~10%) (Figure 37-A). Moreover, around 55% of genes have a peak located within 10kb from their genomic region (Figure 37-B). These results are coherent with what the group of Sun J. and Cvekl.A observed in mouse forebrain and lens [276]. Moreover, Pax6 is known to regulate its own expression and we observed the location of peaks inside P0 promoter region of *pax6b* and inside most of the regulatory regions described in the neighborhood of *pax6b* region (Figure 37-C).

We performed *de novo* DNA sequence motifs detection within the *Pax6b* peaks. The top *de novo* motif identified in consensus peaks (P-value 1e-930) is a 12 bp sequence and analysis of this motif revealed Pax6 known motif as the best match. Indeed, this sequence is highly similar to the *in vivo* motif previously identified in mice lens and forebrain [276] (Figure 38-B) and similar to the Pax6 associated motif on JASPAR database as well (http://jaspar.genereg.net/). This highlights the conservation of Pax6 consensus recognition sequence which is coherent with the strong conservation of PAX6 protein sequence across species (see Introduction 2.3.1). Analysis of the location of the top motif inside peaks showed that it is preferentially located in the center of peaks (Figure 38-C). Interestingly, the following best matches associated to the motif are PAX8, PAX1, PAX9 and PAX5 consensus sequences respectively. These 4 members of the PAX family lack or present a partial homeodomain, supporting the idea that the identified consensus sequence is preferentially associated to the paired domain (PD) [276, 313].

The association of a peak to a neighboring gene is quite challenging. Without additional data, it is difficult to associate ChIP-Seq peaks localized in a genomic region containing multiple close genes. Moreover, some regulatory regions have been described for being

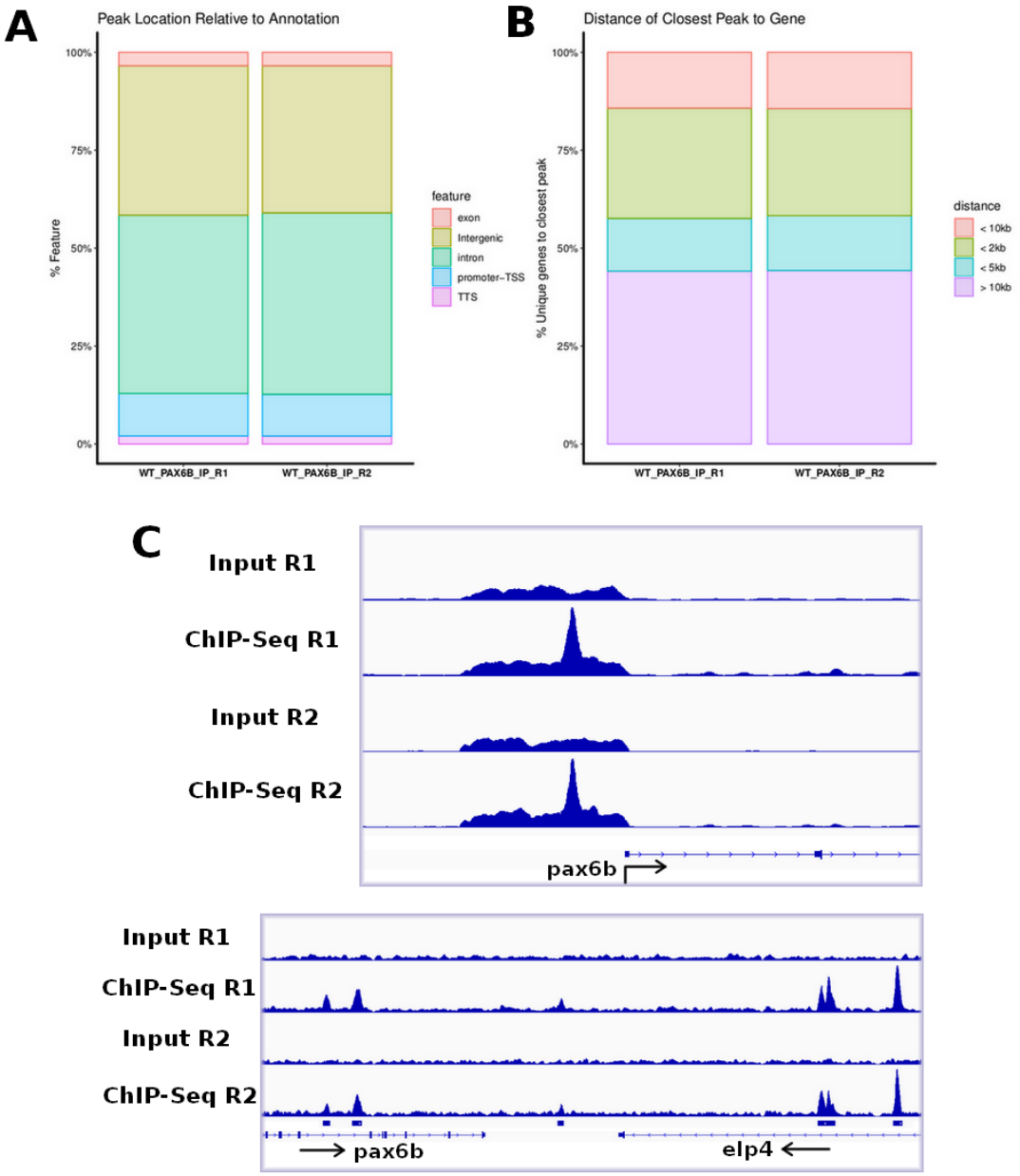


Figure 37: **General analysis of pax6b peaks**. (A) Distribution of peaks between specific genomic regions. (B) Overview of distances between a gene and its closest peak. (C) Identification of pax6b peaks in promoter and regulatory regions of *pax6b*.

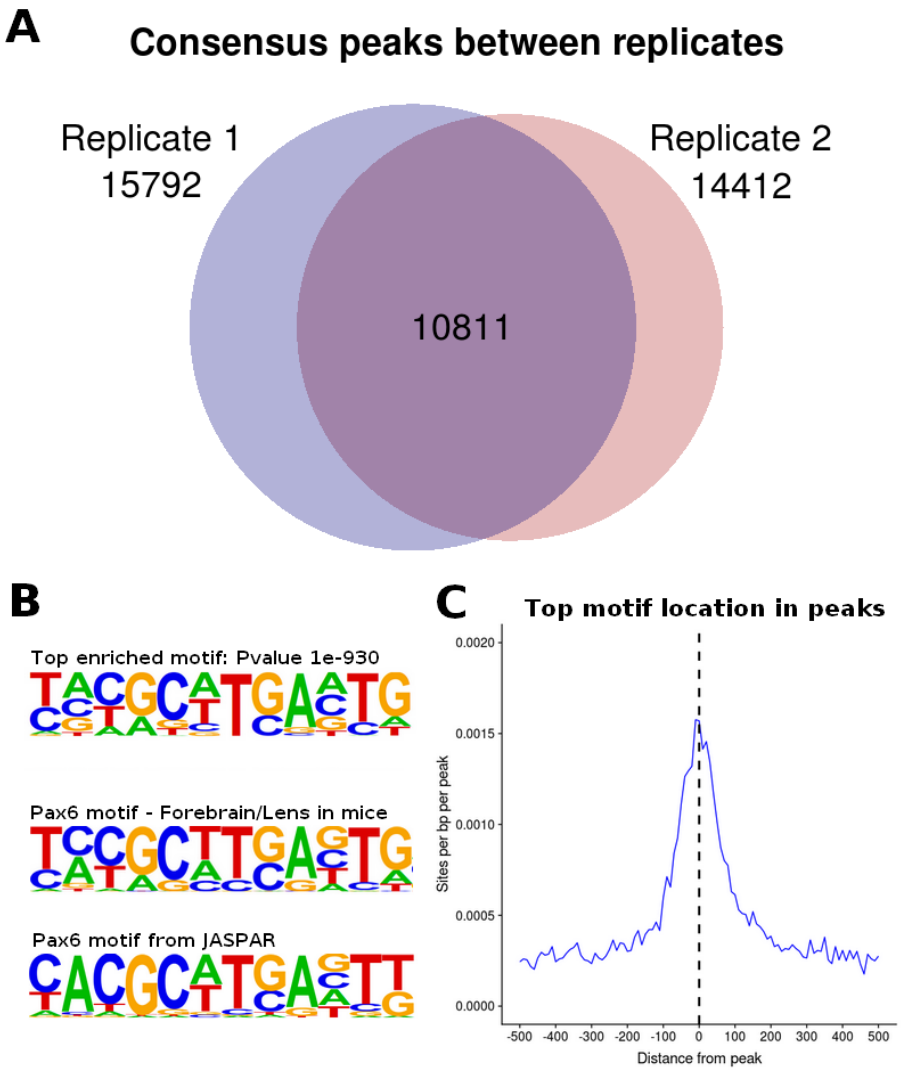


Figure 38: **Validation of peaks**. (A) Identification of consensus peaks between both replicates. (B) Comparison of the top *de novo* motif enriched in consensus peaks with previous *in vitro* and *in vivo* consensus sequence of Pax6. (C) Top motif is enriched in center of the peaks.

very distant of the genes they regulate and evidences have shown that regulatory regions of a gene are sometimes located inside the genomic region of other neighboring genes. We decided to focus on peaks located at less than 10kb of a gene TSS (Transcriptional Start Site) location which reduced our set of identified peaks to 3657 peaks. We linked these selected peaks to their closest gene and we identified 2366 different protein coding genes.

2.2 Pax6b-dependent Gene Regulatory Network (GRN) in endocrine differentiation

As presented in the paper, RNA-Seq datasets have been produced for pancreatic endocrine cells (PECs) isolated from 27 hpf embryos using the transgenic lines *Tg(P0-pax6b:GFP)*, in both wild-type and *pax6b*-mutant (**sa0086** [144]) conditions. The same experiment has been performed for enteroendocrine cells (EECs) using 4 dpf embryos. We compared transcriptomes between wild-type and mutant datasets in order to determine the pax6b regulated genes. Differential expression analyses identified 2824 and 1625 Pax6b-regulated genes in PECs and in EECs, respectively, (FDR<=0.1). Comparison of these two sets of genes reveals that there is a large set of 517 genes which are altered by *pax6b* inactivation in both organs. These results indicate that a part of the Pax6-dependent GRN is thus identical in EECs and PECs.

We decided to identify genes, inside the Pax6 GRN of both tissues, which are linked to a Pax6b peak from the ChIP-seq experiment as these genes are good candidates to be directly regulated. We compared successively our list of 2366 genes linked to pax6b peaks with the list of genes regulated in both tissues, in PECs only and in EECs only. Among the 517 genes regulated in both tissue, 70 (~13.5%) are associated to a close Pax6 peak. Similar ratio of ~15.6% (361/2305) and ~12.2% (136/1116) are observed for genes only regulated in PECs or EECs, respectively (Supplemental table 4). We highlight 11 genes encoding transcription factors bound by Pax6b and regulated in both tissues: *meis1b, rorab, scrt1a, arid5b, pou3f3a, sox4a, atf2, etv5b, hmx4, sp8a* and *neurod1* (Supplemental table 5). Interestingly, mice orthologs of *neurod1* (Neurod1), *pou3f3a* (Pou3f1) and *arid5b* (Arid5b) were previously associated with Pax6 direct regulation in forebrain or lens [276].

We performed gene set enrichment analysis against WikiPathways (WP) and GO Biological Process (GO BP) databases. Considering all 70 targeted genes regulated in both tissues, we identified significative enrichment for MAPK, BMP and FGF signaling pathways and we observed an enrichment of GO terms associated to regulation of multiple metabolic processes (Figure 39). We also looked at targeted genes regulated only in PECs or EECs. For the first group (PECs), we identified many GO terms associated with brain and neuronal development while pathways enrichment highlighted the Delta-Notch Signaling. From EECs list, we observed enrichment for both canonical and noncanonical Wnt pathways. Interestingly, GO terms enrichment analysis of EECs genes additionally highlighted "enteroendocrine cell differentiation", "pancreatic alpha cell differentiation" and "eye morphogenesis" terms (Figure 39).

These results indicate that Pax6b directly controls multiple signaling pathways involved in endocrine differentiation. The similarities between lists of targeted genes with those described in mice forebrain and lens suggest that the Pax6 GRN could be relatively conserved across tissues and species. This is supported by the observation that genes targeted in PECs, EECs or in both are already associated to brain, neuronal or eyes GO terms.

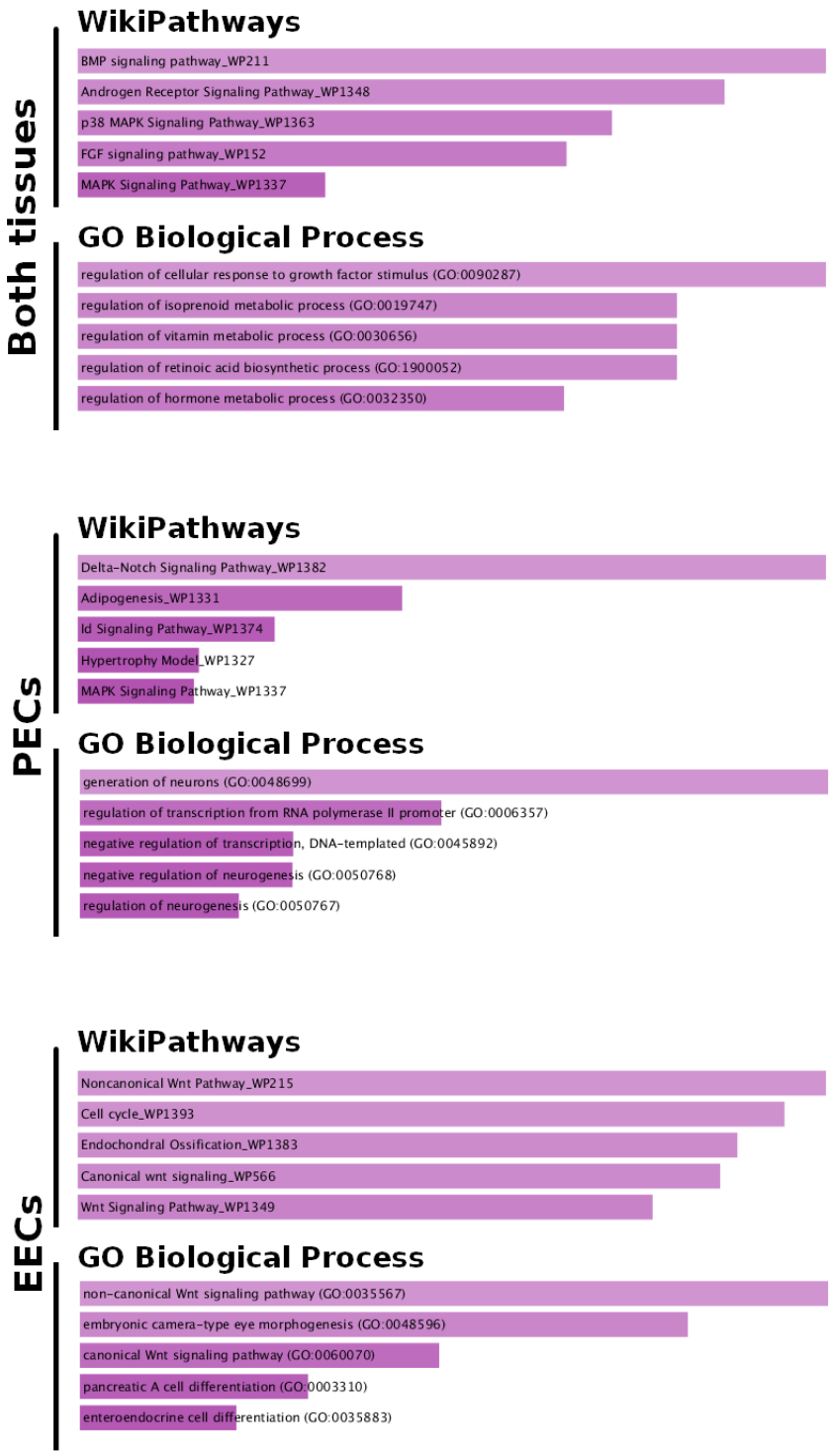


Figure 39: **Gene Set Enrichment Analysis**. WikiPathways and GO BP terms enriched in the 3 lists of targeted genes during endocrine differentiation.

Discussion & Conclusion

1 Transcriptomic signatures of pancreatic cells

In this thesis, we first determined the transcriptomic profiles of pancreatic cell types in zebrafish. We took advantage of using both RNA-seq and scRNA-seq data simultaneously to complete our previous work. The use of scRNA-seq data allowed us to define two supplementary signatures for ghrelin-expressing epsilon cells and for the subpopulation of delta cells expressing specifically *sst1.1* gene. The comparison of all cell type signatures highlighted sets of genes with enriched expression in each cell subtype and validated most of our previous observations [146]. Among these observations, we previously described a set of genes presenting enrichment in endocrine cells compared to ductal and acinar cells and we highlighted that only a small proportion of these genes are specific of an endocrine cell subtype. We performed interspecies comparisons with mammals in order to identify genes with a conserved enriched expression in endocrine cells. Indeed, these genes are expected to be important for endocrine cell functions as their expression has been maintained during evolution of vertebrates. We determined that a set of 251 genes share an endocrine profile through zebrafish, mice and human pancreatic cells. This list contained several transcription factors known for their crucial role in endocrine differentiation such as *runx1t1, isl1, myt1la, fev, neurod1, ascl1a, rfx6, pax6b, insm1b, npas4a, arxa, myt1a, insm1a, lmo1* and *etv1* [146]. We applied a similar approach with genes enriched specifically in alpha and beta cells and we observed that very few genes were conserved between species in this case. Beside the *insulin* gene, only *pdx1* and *gcgra* were conserved for beta cells. Similarly for alpha cells, only a set of 20 genes was found. These results were surprising as they suggest that the identity of the same cell type across species rely on a set of very few conserved genes.

We decided to collect additional data to pursue this analysis and we took advantage of the availibility of scRNA-seq data for pancreatic tissues. As mentioned above, zebrafish scRNA-seq data [147] completed our previous results by filling the remaining gaps. After identification of two supplemental endocrine cell subtype transcriptomes, we reassigned some genes to the corresponding profile. Using human scRNAseq data [134–136], we

obtained transcriptomic signatures for all major pancreatic cell types and we defined all lists of genes with enriched expression in a particular cell type.
For both species, we made a distinction between the enriched genes and the marker genes which are genes presenting a high enrichment in a specific cell type. The strong similarities between the endocrine cell subtypes led to reduced lists of marker genes in both species. Furthermore, the comparison of these lists almost only identified the hormone genes as conserved marker genes in corresponding cell subtype. Thus, this confirmed our previous observations. Nevertheless, we identified *acsl1b*/ACSL1 as an evolutionary conserved marker of epsilon cells which indicates that this enzyme has been conserved in epsilon cell for Acyl-CoA production and making the link with the acylation of ghrelin [297].

We observed an interesting situation with the expression of *pdx1* in zebrafish. Many studies have highlighted the role of PDX1 in pancreatic development, and notably its significance for beta cell differentiation, survival and identity in mice and human [151–154]. In our previous work, we also associated *pdx1* expression to beta cell in zebrafish. However, we obtained the signature of *sst1.1*-expressing delta cell subtype from scRNAseq data and we observed a stronger expression in this cluster. In human scRNA-seq datasets, the expression of PDX1 gene is also detected in delta cells and this observation is coherent with the identification of PDX1 originally as a transcription factor involved in the activation of SST gene [298]. These results suggest that despite its important role in beta cells, PDX1 expression is not restricted to their transcriptome and therefore should not be considered as a marker gene.

Altogether, these results indicate that it is difficult to define a list of genes that highlights specifically an endocrine cell type across species. We evaluated our lists of zebrafish marker genes on human datasets and we could not identify accurately the same cell types. This suggests that looking for specific conserved marker genes is not the best approach and that defining specificity of the different endocrine cell types could also be found in several key genes they express in common. In other words, we suggest that the combination between many endocrine genes expressed in several endocrine cell types could define the specifity of each endocrine cell subtypes. The perspectives to this work would be to

study further the balance between the expression of these endocrine marker genes inside each cell type. Exploring different combination of endocrine genes and evaluating their expression in each endocrine cell subtype could be a more suitable approach. Then, this combination should also be evaluated through interspecies comparison.

1.1 On the importance of datasets

This work relies on the meta-analysis of multiple datasets, meaning the combination of the results of multiple scientific studies. An important aspect to consider is the impact of working with a dataset or another can have on the results. We analyzed all 3 human scRNAseq datasets separately using the same method and we were interested in the overlap between the lists of enriched genes. We observed that around 7% of marker genes of the endocrine subtypes defined for a dataset were validated by the two others datasets. This value rises to around 15% if we consider genes with an enriched expression. Nevertheless, this observation is made under some restrictive conditions as we made the choice to be relatively stringent with our thresholds for the analysis. As we previously described, these 3 scRNAseq datasets were produced by different groups with different SC platforms. Single-cell technology is still in active development and all current platforms suffer from different limitations. Thus, the exact same information is not expected to be covered by the different datasets. We mentioned that from a dataset to another, the distribution of cell populations was heterogeneous and a different amount of cells was used. This experimental variability also contributes to the differences observed in the results.

After we compared the results of each experiment, we decided to proceed through a second method. We integrated all cells together before defining gene enrichment analysis as we expected the aggregation of information to lead to more robust observations. After using batch effect correction of both experimental and technical differences, we highlighted the similarity between datasets and we were able to define lists of enriched genes across platforms for all endocrine cell types.

1.2 On the importance of resources

The field of "Omics" studies has evolved quickly in recent years, and the bioinformatic field followed the same evolution. It implies that the tools and the resources are constantly improving. Online databases allow extracting all sort of available datasets and information while the increasing community of bioinformaticians continually develop tools more suitable for a specific analysis. This work is part of this dynamic and is dependent of different resources like the annotation of the different genomes, the database of orthologous genes across species or the efficiency of genome aligners. A global perspective for this sort of work is to stay up to date with available resources and tools while looking for new datasets in order to increase the robustness of results.

2 The DNA-binding transcription factor PAX6

Most of the results of the second part of this work have been combined in a publication called **"Pancreatic and intestinal endocrine cells share common transcriptomic signatures and gene regulatory networks"** submitted to BMC Biology and currently under reviewing.

In order to understand the mechanisms underlying the Pax6-dependent Gene Regulatory Network in endocrine differentiation, we performed ChIP-seq experiments for Pax6b in zebrafish. We highlighted significative genomic binding sites of Pax6b in whole embryos at 27 hpf and we annotated all peaks with the closest neighboring genes. *De novo* motif detection identified a top motif found in our peaks displaying high similarities with previously identified *in vivo* and *in vitro* Pax6 motifs. The distribution of the genomic locations of Pax6b peaks is coherent with the previous observations [276].

We compared the lists of regulated genes in *pax6b*$^{-/-}$ mutant with the annotated peaks and we identified a set of potential directly targeted genes. The majority of the annotated peaks are not associated to the regulated genes. It suggests that these peaks are inactive or are associated to genes regulated at another stages or in different tissues. Another explanation could be due to the wrong annotation of peaks located in the neighbourhood of several genes. Gene Set Enrichment Analysis on endocrine targeted genes linked

Pax6b to the regulation of many signaling pathways previously described in endocrine differentiation. Furthermore, this analysis also pointed out that several genes directly targeted are associated with brain, neuronal and eye development. This suggests that *pax6* mechanism of regulation could be partially similar between tissues where it is expressed. This is coherent with previous observations of Sun J. and Cvekl A. [276] on the regulatory network of Pax6 during the development of forebrain and lens in mice. Indeed, this study also presented that there is an overlap (around 70%) between peaks identified in the two tissues. Interestingly, we highlighted several transcription factors regulated by Pax6 in endocrine differentiation which they also identified in forebrain or lens development. More recently, the group of Swisa A. and Dor Y. described Pax6 binding-sites in human beta cells close to important endocrine genes such as PDX1, INS, GCG and NKX2.2 [156]. We did not detect any similar peaks close to these genes in our zebrafish ChIP-Seq experiments. It is possible that the amount of pancreatic cells expressing *pax6b* in 27 hpf zebrafish whole embryos is relatively insignificant compared to the mass of *pax6b*-expressing cells found in other tissues such as brain or eyes. It results that the genomic Pax6-binding sites specific of pancreatic cells are not detected in our ChIP-seq dataset. To address this issue, we performed ChIP-seq experiments on adult zebrafish pancreatic tissues. Unfortunately, we did not get any results as we did not retrieve any ChIP-seq signal.

Altogether, these results could indicate that we did not recover pancreatic specific peaks in our datasets and that we may have identified only a part (around 70% based on number of Sun and Cvekl) of Pax6 sites in endocrine cells. It is also reasonable to propose that the significant overlap between the annotated peaks and the genes regulated in endocrine tissues suggests that a part of the Pax6 GRN is similar in the different tissues.

Perspectives for this work would be to successfully identify specific endocrine pancreatic binding-sites of *pax6b* to identify most direct Pax6b target genes and to complete our analysis. This could be performed using most sensitives techniques such as “Cut & Run” [314]. Chromosome conformation capture data, such as Hi-C, would also be of valuable interest in order to efficiently associate a peak to the correct regulated gene(s). In parallel, the comparison of the genomic binding and regulation of Pax6 across species could also

be a strong approach to understand the evolution of Pax6 function and to better resolve its mechanism of action.

Appendices

Supplemental table 1

Set of significative marker genes (pvalue adjusted <= 0.1) identified for alpha, beta and delta cells based on zebrafish bulk RNA-seq data. Genes with an enriched expression in global endocrine cells are also displayed.

Zebrafish RNA-seq Enriched Genes

Gene_Name	Alpha	Beta	Delta	Ductal	Acinar	Cluster	Padj
gpsm1a	5557,4	13,2	25,0	2,2	14,7	Alpha	4,79E-08
zgc:194312	4092,4	99,8	241,9	150,7	282,8	Alpha	3,50E-07
arxa	6780,9	0,3	43,6	7,1	61,2	Alpha	1,01E-05
pcsk5b	2813,9	434,9	399,2	14,2	630,4	Alpha	4,36E-05
epas1b	9982,0	133,5	3323,7	298,6	217,8	Alpha	6,33E-05
etv1	4857,9	50,9	83,9	1,1	9,8	Alpha	1,50E-04
gch2	10977,8	419,3	69,9	30,1	205,3	Alpha	1,70E-04
elovl5	4513,0	288,8	61,0	104,3	76,7	Alpha	2,08E-04
gcga	2238867,1	34692,6	81463,9	97,1	7651,7	Alpha	3,51E-04
pltp	3010,9	26,4	87,4	6,7	53,4	Alpha	3,98E-04
fam107b	8805,1	803,0	1572,8	1086,5	418,8	Alpha	1,16E-03
adora2ab	2062,0	7,3	43,3	6,1	17,6	Alpha	1,49E-03
ptpn11b	3598,3	887,6	548,3	45,0	845,9	Alpha	2,56E-03
tp53i11a	3556,9	50,6	264,4	33,6	27,8	Alpha	2,63E-03
pnoca	249498,5	850,0	13193,0	32,2	978,6	Alpha	3,79E-03
gcgb	768900,3	3866,1	38229,0	57,4	2790,9	Alpha	6,97E-03
dlg1	4270,3	486,1	336,3	478,0	269,5	Alpha	7,23E-03
CABZ01058222.1	1134,4	42,6	182,6	23,3	56,9	Alpha	7,90E-03
pleca	8140,5	817,1	1635,7	1062,6	781,6	Alpha	8,11E-03
sim1a	2050,5	127,7	47,1	103,7	9,8	Alpha	8,82E-03
npb	12456,5	668,1	905,8	28,1	51,0	Alpha	1,05E-02
dzip1	1294,9	0,7	276,2	2,7	2,5	Alpha	1,15E-02
si:dkey-156n14.5	2635,9	534,3	667,9	890,3	620,2	Alpha	1,40E-02
lfng	3223,9	17,9	286,3	127,0	28,2	Alpha	1,43E-02
lingo4b	2138,2	18,8	7,6	1,3	7,4	Alpha	2,16E-02
si:dkey-194e6.1	2110,3	3,2	131,4	57,1	11,4	Alpha	2,54E-02
dusp6	9622,5	921,2	4094,4	1148,6	324,1	Alpha	2,68E-02
fap	1718,0	281,9	43,5	3,8	3,7	Alpha	2,97E-02
slc25a14	6827,6	2938,3	2202,2	1515,6	1927,0	Alpha	3,17E-02
grip1	1174,7	392,5	169,9	5,3	415,5	Alpha	3,28E-02
crhr1	2077,4	47,1	180,6	3,7	4,9	Alpha	3,72E-02
gab1	4709,9	1874,6	642,4	1190,5	917,6	Alpha	3,84E-02
prox1b	2104,4	709,9	677,6	10,5	2,5	Alpha	4,12E-02
si:dkey-172h23.2	1168,9	421,8	579,2	544,1	503,3	Alpha	4,72E-02
robo1	3326,8	1066,6	924,1	51,0	68,3	Alpha	4,95E-02
phkg1a	1504,7	423,0	476,5	17,5	1,6	Alpha	5,07E-02
dpp6a	4028,9	528,7	244,9	0,6	455,1	Alpha	6,17E-02
snrkb	1847,1	194,0	520,8	119,2	372,7	Alpha	6,63E-02
spoplb	1949,6	973,2	668,6	330,1	669,7	Alpha	6,91E-02
mst1	8019,4	1898,9	2,3	439,6	1330,9	Alpha	7,18E-02
gpr142	3098,2	1115,4	120,9	7,3	7,4	Alpha	7,25E-02
pdk2b	14486,1	136,3	700,3	2915,9	1494,3	Alpha	8,32E-02
snap91	5701,1	1209,8	1580,6	164,2	38,0	Alpha	8,34E-02
rasl11a	1056,9	9,6	96,0	7,6	4,9	Alpha	8,39E-02
ptprz1b	1350,9	229,6	116,9	7,8	201,7	Alpha	8,88E-02
pyyb	7428,7	1290,7	622,7	269,0	81,4	Alpha	9,10E-02
mapk14b	2356,7	1043,7	899,5	1212,6	1075,2	Alpha	9,82E-02
srpx	265,8	2321,5	24,5	104,4	4,9	Beta	2,53E-13
ndufa4l2a	245,9	24138,4	104,2	125,3	12,7	Beta	7,87E-10
tmeff2a	204,7	2909,9	32,5	2,2	0,9	Beta	4,34E-09
esr1	393,1	5961,2	40,8	45,8	17,9	Beta	1,39E-07
ins	20304,6	3427305,3	17150,6	11323,1	3799,2	Beta	1,60E-07

fgfr1b	149,6	2173,2	40,5	29,7	0,0	Beta	8,98E-07
nkx6.2	32,3	2512,3	4,6	8,3	0,0	Beta	3,34E-06
ngfra	225,4	2374,8	37,1	8,3	5,8	Beta	9,30E-06
ca12	144,8	4171,0	32,2	50,9	6,0	Beta	1,76E-05
rrad	65,2	1884,9	43,2	64,8	67,1	Beta	4,16E-05
rasd4	447,1	3279,9	53,2	19,2	0,0	Beta	1,14E-04
CABZ01085139.1	58,7	1741,6	90,7	22,0	48,2	Beta	1,34E-04
tagln3b	947,4	11943,9	1122,3	691,6	507,8	Beta	3,87E-04
kcnj2a	244,4	7506,3	373,9	5,0	31,4	Beta	4,07E-04
wif1	1191,0	9273,0	206,2	1346,5	108,4	Beta	4,25E-04
prkcbb	43,9	2682,2	204,5	31,5	22,7	Beta	5,61E-04
zgc:154058	404,5	2323,5	384,2	109,6	176,2	Beta	6,70E-04
grapb	254,5	1110,0	3,2	180,5	1,6	Beta	7,88E-04
ACVR1C	110,1	2611,8	385,3	11,5	12,5	Beta	2,12E-03
otc	92,3	3031,0	43,7	1,4	46,9	Beta	2,80E-03
ucn3l	41,7	5036,6	145,9	6,4	0,0	Beta	3,15E-03
SLC5A10	11283,3	40227,2	438,3	584,6	231,3	Beta	4,17E-03
dpp6b	129,5	2829,1	205,5	8,1	25,4	Beta	4,24E-03
ca7	62,9	6931,1	20,2	40,2	129,1	Beta	4,28E-03
pcdh10a	1435,6	17274,0	3142,8	24,4	4,6	Beta	4,68E-03
ace	28,0	1717,2	44,6	4,9	27,5	Beta	5,97E-03
fstl1b	180,6	6201,3	49,7	377,7	225,3	Beta	6,06E-03
ppdpfb	120746,7	822689,0	29145,1	18155,7	36335,1	Beta	6,15E-03
nptna	750,7	5272,8	97,1	39,1	5,7	Beta	6,65E-03
idh3a	2471,6	4220,3	2464,7	2754,2	1893,5	Beta	7,75E-03
phyhiplb	587,6	2192,5	650,3	1,4	5,8	Beta	8,99E-03
hdac9b	2478,5	7069,7	1939,4	679,5	17,6	Beta	1,04E-02
rnd1b	131,8	1114,4	3,1	2,0	1,3	Beta	1,10E-02
nkx3.2	171,3	1049,0	1,1	74,8	0,0	Beta	1,55E-02
ergic1	322,9	1930,3	377,0	681,8	554,5	Beta	1,56E-02
man2c1	375,4	2246,7	192,0	159,1	135,3	Beta	1,63E-02
g6pcb	2023,6	19787,6	232,7	63,3	20,3	Beta	1,95E-02
stmn4	58,3	1333,5	1,6	62,4	23,0	Beta	2,16E-02
tmtc1	397,1	2168,0	340,1	162,8	90,7	Beta	2,41E-02
slc38a11	72,9	1332,3	217,4	0,3	0,9	Beta	3,30E-02
btc	1650,4	5879,5	1719,0	507,1	558,6	Beta	3,70E-02
nudcd2	355,2	1410,9	382,0	543,8	367,1	Beta	4,07E-02
klb	439,8	2173,9	328,5	12,9	12,9	Beta	4,31E-02
tmsb2	353,3	1032,0	373,5	5,9	0,0	Beta	4,80E-02
sec23a	1410,8	2398,3	961,3	944,9	1418,8	Beta	5,41E-02
BX927333.1	270,7	1115,7	335,4	5,4	2,5	Beta	5,74E-02
prkdc	338,3	2537,5	211,9	57,8	86,6	Beta	5,92E-02
emc10	3632,1	7598,2	2438,1	2231,2	3355,7	Beta	6,29E-02
tmem163b	86,5	1505,2	380,1	5,9	1,6	Beta	6,45E-02
calua	948,5	3044,0	1234,9	603,4	1228,2	Beta	6,83E-02
si:dkey-57n24.6	543,9	1016,8	346,7	253,6	329,4	Beta	6,83E-02
osbpl5	471,7	1317,0	430,5	279,5	35,9	Beta	7,45E-02
clstn3	203,6	1631,1	82,7	596,7	140,5	Beta	7,46E-02
golga7ba	508,7	1478,1	686,7	11,2	0,0	Beta	7,80E-02
nansa	1236,2	3144,0	1010,2	1422,3	1558,4	Beta	8,52E-02
crtac1a	459,3	1872,3	253,7	15,5	1,6	Beta	8,59E-02
pcsk2	4894,1	28108,2	164,9	65,9	25,3	Beta	9,46E-02
LO017791.1	29,0	0,4	1293,0	0,3	0,0	Delta	5,59E-27

UTS2R	0,7	0,1	1009,3	1,0	0,0	Delta	3,62E-25
arl3l2	16,4	5,5	2277,2	9,2	5,8	Delta	4,47E-21
lamc2	90,7	8,5	8680,5	12,6	109,2	Delta	7,16E-18
mkxa	58,9	0,2	3493,7	1,3	0,0	Delta	2,51E-15
LO018309.1	17,5	9,9	8669,8	5,2	4,1	Delta	5,72E-13
cdx1b	131,1	5,6	10738,8	12,4	11,1	Delta	1,38E-12
map2k6	2770,2	2450,1	45161,8	998,3	755,2	Delta	2,11E-12
TMEM179 (1 of many)	224,5	77,9	4235,6	0,0	0,0	Delta	5,50E-12
map6b	66,7	45,1	1522,2	0,3	0,0	Delta	2,23E-11
sgcd	17,1	0,3	2075,4	0,5	0,0	Delta	4,05E-11
ppp1caa	3881,9	4661,3	24959,5	4981,2	2208,3	Delta	3,60E-10
stxbp6	2020,0	608,3	18576,7	1274,5	1552,7	Delta	4,45E-10
marcksl1b	132,1	115,8	13629,0	271,9	118,8	Delta	6,34E-10
si:dkey-206p8.1	230,8	130,6	5843,1	1,5	1,3	Delta	7,77E-10
rnd2	154,4	0,0	1261,8	0,0	0,0	Delta	1,39E-09
anxa11a	2226,5	1867,9	27540,2	3172,2	2991,7	Delta	5,81E-09
ddc	18,8	0,5	2759,8	0,8	11,0	Delta	6,21E-09
sall4	57,3	28,6	1444,7	131,1	109,9	Delta	1,05E-08
napbb	587,9	421,5	3126,0	9,6	5,7	Delta	1,74E-08
tub	122,3	2,1	1841,7	5,1	13,8	Delta	2,86E-08
sst2	24375,0	8862,7	4878574,5	359,9	4261,7	Delta	3,20E-08
nrcama	168,4	107,9	3171,1	6,7	0,0	Delta	7,12E-08
mdkb	106,4	41,4	10653,0	310,3	77,8	Delta	9,51E-08
aebp1	459,2	18,1	42924,0	4,9	77,7	Delta	1,99E-07
chgb	94,9	94,2	1790,3	139,8	12,3	Delta	2,15E-07
ttc6	31,5	0,5	1198,2	34,2	31,9	Delta	2,78E-07
adra1d	29,4	8,0	3887,1	2,9	2,5	Delta	6,04E-07
gulp1b	104,8	47,4	11926,3	129,5	56,1	Delta	9,20E-07
kcnt1	43,7	5,2	2402,1	0,1	0,0	Delta	1,36E-06
ifitm5	14,1	3,8	1122,5	20,0	22,0	Delta	1,42E-06
spegb	54,1	18,0	6694,5	9,7	30,6	Delta	2,01E-06
rprma	55,8	24,6	1095,5	1,0	0,9	Delta	2,08E-06
pear1	18,3	1,8	1182,0	78,0	5,4	Delta	2,29E-06
mmp2	26,1	0,4	2169,4	21,9	19,2	Delta	3,53E-06
avpr2l	13,7	2,4	2067,8	0,7	0,0	Delta	3,54E-06
phkg2	650,3	603,5	3117,3	86,4	250,1	Delta	5,03E-06
gng3	83,0	225,3	3097,6	5,9	0,0	Delta	8,29E-06
syt10	85,7	0,0	2120,2	20,5	0,0	Delta	9,15E-06
mertka	3,5	1,3	2836,6	224,9	8,0	Delta	1,10E-05
map3k15	329,8	10,4	18042,8	130,8	52,6	Delta	1,15E-05
sst1.2	17742,4	8187,0	3429815,4	392,7	3916,2	Delta	1,30E-05
ptger2a	116,8	5,8	8558,1	0,8	1,6	Delta	1,71E-05
srpx2	181,4	439,9	2927,6	82,4	0,0	Delta	2,62E-05
bmp7b	12,1	1,8	1535,1	22,3	6,5	Delta	2,65E-05
cldni	165,3	9,6	17568,3	3,3	7,4	Delta	2,72E-05
nmba	338,8	90,7	19070,6	11,0	8,5	Delta	2,81E-05
SBSPON	3,5	3,9	1088,0	63,6	0,0	Delta	3,33E-05
rxfp2a	128,7	221,7	2592,4	265,2	432,6	Delta	5,33E-05
mapk14a	3232,3	1812,2	16045,0	3457,5	960,6	Delta	6,17E-05
zgc:195023	303,7	55,5	44732,6	2,6	82,8	Delta	6,94E-05
ugp2b	3358,8	1697,1	10913,2	1029,3	927,1	Delta	8,87E-05
aqp4	52,6	3,2	3797,7	0,3	0,0	Delta	9,14E-05
carhsp1	600,8	121,8	15840,7	128,1	1791,0	Delta	9,20E-05

st3gal3a	327,9	250,3	1303,3	209,2	202,9	Delta	9,35E-05
pfas	280,7	469,6	5069,7	193,6	410,3	Delta	1,02E-04
sema3fa	78,8	6,9	5020,6	8,2	109,4	Delta	1,07E-04
GALNTL6	10,1	0,8	1182,9	1,5	28,4	Delta	1,24E-04
tbx2b	99,2	30,4	6639,2	4,0	15,9	Delta	1,31E-04
glra1	28,4	2,2	1554,0	51,3	1,6	Delta	1,44E-04
si:ch211-195b15.8	635,2	143,8	11930,9	910,3	131,3	Delta	1,62E-04
fabp6	72,5	58,6	12285,3	16,4	24,6	Delta	1,86E-04
inpp5kb	123,2	6,5	2195,7	0,3	154,2	Delta	2,11E-04
cpamd8	236,0	41,9	7657,6	8,3	1,6	Delta	2,50E-04
ank2a	38,7	39,6	1731,8	2,0	24,3	Delta	2,54E-04
srd5a2b	114,9	72,1	5788,9	106,8	9,8	Delta	3,45E-04
CABZ01088484.1	33,5	7,9	1862,9	34,1	32,0	Delta	4,12E-04
agtr1b	74,2	55,2	1509,4	0,0	9,0	Delta	4,21E-04
cdh11	69,5	6,2	4528,1	2,5	58,6	Delta	4,63E-04
si:dkey-166d12.2	15,4	0,5	1718,6	134,7	7,7	Delta	4,91E-04
plppr5a	20,9	0,6	1830,3	0,3	90,4	Delta	5,35E-04
wnk4b	279,5	615,2	4477,1	9,6	6,6	Delta	5,41E-04
prnpb	39,8	14,2	1001,3	1,1	0,0	Delta	5,55E-04
prkg1b	55,9	36,8	2940,3	202,6	52,4	Delta	5,59E-04
CABZ01068367.1	15,6	0,1	1308,2	209,1	164,0	Delta	7,09E-04
ca6	24,5	3,8	2267,6	5,6	141,1	Delta	8,34E-04
CU019662.1	26,5	2,9	1790,8	0,5	14,6	Delta	9,45E-04
lrrc39	89,6	56,1	1251,0	30,3	45,2	Delta	1,03E-03
amd1	5764,3	3267,4	25930,8	4961,3	1691,1	Delta	1,09E-03
zgc:100906	274,1	322,6	3538,5	72,6	96,3	Delta	1,15E-03
slc26a3.2	130,8	97,5	12841,0	898,0	38,8	Delta	1,35E-03
ghrl	5476,0	487,3	123262,0	6318,3	6749,1	Delta	1,38E-03
ncf1	159,6	91,3	3786,1	269,1	403,7	Delta	1,70E-03
itga11b	38,8	6,9	1227,2	3,7	161,4	Delta	1,77E-03
smoc2	15,3	0,1	2269,9	76,1	9,7	Delta	1,94E-03
serpini1	619,0	550,9	9917,2	9,6	69,4	Delta	1,99E-03
zgc:153615	103,6	704,9	1957,4	0,5	0,0	Delta	2,08E-03
nrbp2a	62,2	31,8	1690,1	10,1	30,8	Delta	2,18E-03
HHEX	92,1	55,0	8850,1	626,0	62,4	Delta	2,20E-03
srgap3	362,9	725,9	1969,6	53,1	241,6	Delta	2,97E-03
trpm3	36,0	0,0	1124,7	5,4	41,7	Delta	3,12E-03
ptger1b	17,2	5,3	2830,7	0,4	0,0	Delta	3,27E-03
kctd12.2	1380,0	181,2	66319,7	5043,6	257,2	Delta	3,32E-03
ak5	107,5	417,8	2183,5	13,4	1,7	Delta	3,35E-03
atp1a1b	69,6	4,5	12517,8	6,8	528,7	Delta	3,48E-03
tspan4b	17,0	0,0	2233,2	7,1	92,4	Delta	3,55E-03
irs1	108,2	202,6	1498,2	228,3	236,6	Delta	3,58E-03
slc7a14a	338,2	309,3	3888,2	21,6	119,5	Delta	3,72E-03
slc7a7	965,2	94,2	17935,0	26,7	41,2	Delta	3,75E-03
zgc:198419	535,7	566,5	2222,6	193,0	143,9	Delta	3,83E-03
si:dkey-280e21.3	262,4	106,6	3753,5	14,7	0,0	Delta	4,22E-03
slmapb	670,8	100,1	6707,6	15,0	35,5	Delta	4,26E-03
pcdh9	184,9	161,8	2475,5	1,9	0,0	Delta	4,35E-03
mrap2a	261,9	2,9	7126,1	1,7	18,4	Delta	4,59E-03
slc8a1b	25,1	171,4	1243,2	1,3	0,0	Delta	4,98E-03
prss23	16,4	6,7	1240,2	6,7	182,1	Delta	5,31E-03
npy8br	135,9	1,0	2776,7	26,2	45,7	Delta	5,91E-03

oaz2a	2677,6	1530,2	17597,6	2098,5	1105,5	Delta	6,59E-03
FO904977.1	25,1	32,9	1244,1	2,6	2,5	Delta	6,69E-03
olfm2a	447,8	24,9	10411,2	4,6	4,9	Delta	7,35E-03
fgfr1a	375,2	101,5	4338,5	9,7	3,3	Delta	8,53E-03
adcyap1r1a	56,2	108,5	1486,3	10,1	18,8	Delta	9,06E-03
pitpnaa	441,2	1212,2	20193,7	2432,1	1178,7	Delta	9,87E-03
TMEM8B	161,5	125,9	1041,8	2,1	0,9	Delta	1,01E-02
si:dkey-85k7.7	117,0	107,0	3517,3	323,5	256,9	Delta	1,03E-02
trib2	1007,1	743,0	4684,4	91,2	923,7	Delta	1,06E-02
tmed3	328,6	384,3	1806,3	115,1	522,3	Delta	1,11E-02
yjefn3	135,0	3,2	3089,0	1,7	9,0	Delta	1,16E-02
mpped2a	587,5	402,9	1217,8	429,1	397,4	Delta	1,20E-02
tmcc3	216,5	64,3	1922,8	134,5	381,7	Delta	1,37E-02
DUT	364,3	678,9	2162,1	163,4	20,1	Delta	1,37E-02
slc35g2b	71,3	30,0	3062,0	18,8	0,0	Delta	1,44E-02
edil3a	72,1	4,9	4802,0	651,8	105,6	Delta	1,46E-02
rasgef1ba	1725,6	1959,3	11170,9	2997,2	1901,8	Delta	1,49E-02
kcnk3b	850,2	1092,3	15042,5	13,8	41,7	Delta	1,59E-02
apaf1	382,8	283,5	1230,9	264,0	490,9	Delta	1,66E-02
ppp6c	2127,9	2715,0	7828,9	2852,0	1391,8	Delta	1,67E-02
csrnp2	199,5	182,4	1549,7	42,5	2,9	Delta	1,69E-02
stom	347,4	349,5	2017,2	337,0	73,1	Delta	1,86E-02
map1ab	1341,8	1003,6	12744,2	14,6	546,3	Delta	1,97E-02
hmx4	22,9	2,5	1405,6	6,4	84,0	Delta	2,00E-02
hopx	428,1	699,5	3192,8	437,1	307,9	Delta	2,06E-02
cyp3a65	477,8	208,1	1226,8	18,6	0,0	Delta	2,13E-02
nhsb	303,3	115,3	2325,4	25,3	6,2	Delta	2,27E-02
brd1b	1458,7	1532,4	2729,0	702,2	1308,7	Delta	2,51E-02
stk3	336,9	572,6	1608,2	691,4	561,4	Delta	2,77E-02
iba57	263,8	280,6	1141,2	382,9	441,0	Delta	2,84E-02
loxl2a	100,0	49,3	1842,4	1,6	12,2	Delta	2,86E-02
skib	451,7	337,9	1250,9	159,0	1,6	Delta	2,91E-02
lancl2	720,9	578,9	1751,3	223,1	243,1	Delta	2,93E-02
tspan7	2505,4	2323,5	6459,3	1911,0	368,8	Delta	3,24E-02
si:dkey-250k15.7	35,1	10,1	1400,3	418,9	103,8	Delta	3,48E-02
alcamb	19090,5	11010,3	66011,6	1536,9	484,4	Delta	3,62E-02
CABZ01079192.1	147,4	28,6	8129,1	480,4	129,6	Delta	3,63E-02
tenm3	932,0	982,8	2235,2	58,2	18,3	Delta	3,98E-02
nr0b1	49,2	186,9	1416,2	26,6	21,2	Delta	4,06E-02
tshz3a	63,5	0,6	2713,5	4,8	265,4	Delta	4,08E-02
pygma	182,5	52,7	3034,7	13,3	129,7	Delta	4,14E-02
spock3	447,0	1794,8	7789,7	147,4	2,5	Delta	4,80E-02
cdh13	423,3	1,3	1782,1	4,3	0,0	Delta	5,00E-02
maptb	203,1	599,1	3496,4	64,8	22,6	Delta	5,03E-02
prkag1	1299,8	1178,3	3343,6	1098,9	793,2	Delta	5,21E-02
adgrv1	142,8	241,1	1193,9	14,1	64,8	Delta	5,23E-02
bsnb	618,2	683,0	1665,3	26,2	0,0	Delta	5,24E-02
ckbb	1954,2	3046,3	18489,4	129,4	472,9	Delta	5,31E-02
rnf19a	1285,0	702,2	3395,4	1160,6	965,8	Delta	5,39E-02
gldc	964,4	21,6	3384,0	313,4	679,2	Delta	5,97E-02
rab30	293,5	272,1	1913,2	75,3	74,2	Delta	6,02E-02
tfcp2	1586,3	1854,5	5536,5	1370,8	357,8	Delta	6,21E-02
lamb2	535,9	0,6	3843,1	11,2	9,7	Delta	6,46E-02

slc26a5	1945,1	495,9	5188,0	1192,5	1822,0	Delta	6,48E-02
casr	263,9	237,5	1153,2	7,7	2,6	Delta	6,52E-02
scube2	692,4	267,5	3580,2	759,5	47,0	Delta	6,80E-02
zgc:55733	475,5	333,2	1560,0	671,1	327,8	Delta	6,84E-02
gpr4	297,6	13,4	2091,9	1,0	2,5	Delta	6,87E-02
p3h4	171,7	20,9	2299,6	8,6	411,5	Delta	7,04E-02
stx2a	57,9	44,5	1780,5	703,8	14,0	Delta	7,08E-02
nucb2b	6185,7	7984,4	18667,0	351,0	6027,1	Delta	7,20E-02
stmn1b	2273,1	486,9	14825,5	44,8	275,1	Delta	7,47E-02
cplx2	369,5	742,7	2515,6	19,0	1,6	Delta	7,93E-02
mfge8a	289,7	300,1	1085,9	108,8	317,8	Delta	8,88E-02
hepacam2	365,1	10,3	4919,7	53,4	66,5	Delta	8,94E-02
aatka	1154,2	811,1	404,6	17,2	2,5	Endocrine	-
abca2	1040,4	1084,3	530,9	124,1	9,8	Endocrine	-
abcc5	13528,1	3901,8	4878,6	425,3	626,9	Endocrine	-
abcc8	3289,3	7471,2	5124,6	46,8	21,2	Endocrine	-
abhd15a	1669,7	10188,0	28,7	173,2	131,3	Endocrine	-
adam19b	1467,1	147,8	0,1	4,1	20,3	Endocrine	-
adgrb1a	948,3	905,8	2645,3	12,0	85,5	Endocrine	-
adgrb1b	2370,0	854,2	3024,0	14,8	29,6	Endocrine	-
adgrb3	1140,0	858,5	1284,1	10,9	15,9	Endocrine	-
adm2a	100664,2	59441,0	29395,1	604,5	2342,6	Endocrine	-
ajap1	342,4	1206,4	886,0	4,5	11,7	Endocrine	-
ak5l	10777,8	3521,9	308,2	25,6	39,2	Endocrine	-
aldocb	34716,1	36156,2	18672,4	207,9	578,7	Endocrine	-
amigo1	573,9	667,2	2137,0	0,8	0,0	Endocrine	-
amigo3	2787,3	3566,0	392,4	37,1	20,3	Endocrine	-
amph	2908,2	2325,1	2313,1	9,1	209,9	Endocrine	-
angptl3	6709,1	1161,7	35,1	9,3	19,6	Endocrine	-
ank2b	4632,1	1294,8	1190,2	148,6	71,0	Endocrine	-
ank3b	1446,7	1205,8	1357,3	6,5	37,6	Endocrine	-
ankrd13d	1313,9	1304,2	1010,4	5,7	41,7	Endocrine	-
anks1b	563,2	133,9	1117,8	14,7	2,6	Endocrine	-
ano8a	1074,7	386,6	351,6	1,2	11,4	Endocrine	-
anxa13l	175,7	772,1	1125,3	13,7	0,0	Endocrine	-
anxa5a	1046,2	2061,2	1547,7	208,2	24,2	Endocrine	-
ap1s2	2680,2	5364,9	3563,5	167,0	169,7	Endocrine	-
apba1b	1967,4	1156,5	1490,2	25,2	156,6	Endocrine	-
apof	10216,8	1501,3	4791,8	7,6	27,0	Endocrine	-
arf3a	386,7	754,0	1002,0	3,8	124,8	Endocrine	-
arhgef12a	2973,6	2762,5	1894,1	186,4	161,9	Endocrine	-
arhgef9a	1284,0	384,9	605,6	2,9	28,1	Endocrine	-
arvcfb	3207,4	968,8	2082,4	133,9	49,0	Endocrine	-
astn1	3437,2	2486,5	3600,0	13,5	85,9	Endocrine	-
atcaya	2034,2	3067,0	5429,9	24,7	2,5	Endocrine	-
atp10a	1422,9	1304,2	1290,9	162,2	35,7	Endocrine	-
atp1a3a	16861,1	10052,4	4426,7	116,2	196,5	Endocrine	-
atp1b2a	2266,3	1183,8	1269,2	75,2	23,6	Endocrine	-
atp2b1b	940,8	500,4	1102,3	8,3	20,3	Endocrine	-
atp2b3a	2308,8	1762,8	795,1	8,6	40,5	Endocrine	-
atp6ap1a	1676,5	2395,0	2070,6	31,8	2,5	Endocrine	-
atp6v0a1b	699,2	710,0	1353,3	8,4	30,6	Endocrine	-
atp6v1c1b	1057,4	1204,5	1046,5	96,0	114,9	Endocrine	-

ATP8A1	1179,2	192,1	238,1	1,1	30,1	Endocrine	-
avpr1aa	1272,7	2175,7	1137,4	0,3	0,0	Endocrine	-
bean1	1332,0	1522,3	854,6	5,8	6,2	Endocrine	-
bin1a	1009,6	726,8	376,0	1,2	44,8	Endocrine	-
c2cd4a	12840,7	5796,8	18816,3	936,1	137,6	Endocrine	-
ca16b	555,4	1596,5	389,8	52,6	0,9	Endocrine	-
CABZ01038494.1	506,3	1028,6	320,3	11,1	19,4	Endocrine	-
CABZ01073265.1	22760,9	23448,1	15864,5	457,1	341,1	Endocrine	-
CABZ01102528.1	535,4	890,5	1100,9	80,4	14,5	Endocrine	-
cacna1c	1512,4	576,5	700,1	3,8	80,2	Endocrine	-
cacna1da	1373,5	447,8	227,5	3,5	10,6	Endocrine	-
cacna2d2a	1450,6	1057,8	528,2	9,0	3,3	Endocrine	-
cacnb4b	2838,5	1029,5	820,9	7,3	9,8	Endocrine	-
cadm1a	1158,9	1487,3	262,5	3,1	2,5	Endocrine	-
cadm2a	2349,3	6898,6	3594,7	226,5	216,8	Endocrine	-
caly	518,8	595,9	1742,5	1,4	33,9	Endocrine	-
camk1da	663,8	870,2	1196,5	35,6	13,9	Endocrine	-
camk2n1a	8677,2	2551,8	10097,1	30,9	68,5	Endocrine	-
camkvb	1802,1	1,2	2461,8	5,1	31,0	Endocrine	-
camta1b	1438,2	655,4	996,7	31,3	2,5	Endocrine	-
ccdc85al	1281,8	1165,4	92,2	31,9	2,5	Endocrine	-
cdh18a	43,5	673,5	1316,7	3,7	0,0	Endocrine	-
cdh4	178,3	341,5	1291,6	3,9	3,2	Endocrine	-
cdh6	1721,4	2321,8	676,3	11,5	7,4	Endocrine	-
CDHR2	1568,5	565,8	1778,1	7,8	4,9	Endocrine	-
cdk5r2b	2393,2	7,4	3343,5	2,3	4,9	Endocrine	-
celf3a	1615,4	1609,2	1456,4	32,7	21,2	Endocrine	-
celf5a	1009,9	590,2	2041,1	2,2	17,8	Endocrine	-
cers6	2720,2	1461,8	1703,3	15,8	14,7	Endocrine	-
chd5	489,6	1217,6	1407,4	11,4	17,0	Endocrine	-
chga	27414,8	10010,4	3946,7	89,8	116,6	Endocrine	-
chrnb3b	2232,1	1071,4	3226,1	1,1	9,0	Endocrine	-
clip3	3818,2	2815,9	2689,4	5,0	21,4	Endocrine	-
cntn4	249,0	1257,3	19,3	1,2	3,3	Endocrine	-
cntn5	1721,2	2627,9	47,3	1,0	15,4	Endocrine	-
cntnap2a	5988,0	9779,0	369,9	50,2	110,8	Endocrine	-
coro1cb	671,9	735,2	1923,3	4,4	2,5	Endocrine	-
cpe	68496,0	87167,6	42017,2	1090,1	862,3	Endocrine	-
cplx2l	4447,9	3033,4	2086,3	305,4	44,9	Endocrine	-
creb3l1	1494,4	1367,6	2610,4	39,8	47,4	Endocrine	-
CSDC2 (1 of many)	1040,0	368,2	384,3	17,9	84,2	Endocrine	-
ctnna2	1813,3	1577,7	1843,3	3,7	18,4	Endocrine	-
ctnnal1	904,6	1144,4	850,8	110,6	36,5	Endocrine	-
ctnnd2b	2100,7	1280,4	868,9	14,4	7,4	Endocrine	-
CU855878.1	4511,8	1736,8	928,7	162,7	6,2	Endocrine	-
cyb561	2212,9	3528,3	2432,1	15,8	34,8	Endocrine	-
cyfip2	2601,6	2147,0	2908,0	38,7	17,5	Endocrine	-
cyp27c1	1232,4	1562,9	2,4	52,8	6,5	Endocrine	-
cyr61l2	2015,1	1086,2	9,9	44,3	7,4	Endocrine	-
cyth1b	4083,9	3680,9	5697,0	232,1	299,1	Endocrine	-
dab2ipa	1078,0	979,7	431,3	21,9	11,7	Endocrine	-
dachd	1239,6	908,2	239,8	93,8	2,5	Endocrine	-
dclk1a	1981,9	2109,1	1692,0	12,9	31,9	Endocrine	-

desi1a	1389,6	2092,7	1567,3	260,3	60,8	Endocrine	-
dgkaa	2528,4	5201,9	30,7	108,5	134,0	Endocrine	-
dgkh	2401,2	865,1	132,1	47,3	79,2	Endocrine	-
disp2	8471,7	10007,9	6914,3	27,6	89,4	Endocrine	-
dkk3b	115090,2	8093,6	59550,0	2329,0	648,6	Endocrine	-
dnajc5aa	1133,5	1086,3	1192,3	21,8	155,9	Endocrine	-
doc2d	1117,2	552,7	51,7	0,6	4,1	Endocrine	-
dock9b	456,8	1250,8	1015,7	8,5	3,3	Endocrine	-
dpysl2b	1165,9	2044,6	1315,2	2,6	11,3	Endocrine	-
dpysl3	504,3	1065,3	334,5	14,7	56,0	Endocrine	-
dscama	6875,3	3226,2	193,7	11,4	109,3	Endocrine	-
ECE2	1268,4	426,7	1039,0	46,8	50,3	Endocrine	-
efnb2a	1825,5	3774,1	2354,4	270,7	258,3	Endocrine	-
egr4	38275,4	35511,9	69064,9	341,0	706,0	Endocrine	-
ehd4	1418,4	922,5	1986,5	59,4	225,7	Endocrine	-
ek1	2133,6	864,6	920,1	128,5	67,0	Endocrine	-
elavl4	2911,8	1668,5	2096,1	17,0	4,9	Endocrine	-
elmod1	696,5	306,1	1182,4	2,6	13,0	Endocrine	-
elovl1a	7621,2	5325,7	7017,2	1079,9	32,3	Endocrine	-
eno1a	8408,6	19506,0	11738,2	923,7	1086,1	Endocrine	-
epdr1	2034,6	3986,7	1535,4	23,8	89,6	Endocrine	-
epha4b	1899,8	2092,3	1519,8	16,3	4,9	Endocrine	-
etv4	111,8	3,5	1269,5	4,3	70,9	Endocrine	-
etv5a	6424,3	3120,3	1950,4	392,6	272,9	Endocrine	-
etv5b	7712,1	1804,9	2786,1	502,6	256,1	Endocrine	-
eva1a	1223,7	1666,9	1848,4	30,6	5,8	Endocrine	-
fam131c	965,4	1426,4	33,6	12,6	11,7	Endocrine	-
fam19a5a	1110,2	81,5	305,6	1,1	0,0	Endocrine	-
fam212aa	149,8	988,4	4348,6	51,2	21,2	Endocrine	-
fam43a	2313,2	1061,0	14,0	61,1	26,3	Endocrine	-
FAM46A	2641,5	1244,6	1536,6	13,8	13,9	Endocrine	-
fam49bb	8810,7	7522,8	6157,3	333,5	277,4	Endocrine	-
fam69c	943,0	1560,7	46,6	18,5	0,0	Endocrine	-
fam83fa	1646,5	702,2	266,8	154,7	5,7	Endocrine	-
FBLN1	1349,3	69,8	658,2	1,5	7,6	Endocrine	-
fev	16239,0	687,0	8412,6	58,9	198,4	Endocrine	-
fez1	2691,0	3292,5	1803,9	8,4	22,0	Endocrine	-
flot1b	3979,5	4629,3	1322,9	108,5	53,0	Endocrine	-
flrt3	5208,5	906,4	2826,3	132,4	304,1	Endocrine	-
fndc4b	1885,5	1156,7	39,3	15,4	4,2	Endocrine	-
fndc5b	2562,5	1589,4	1132,1	1,8	45,9	Endocrine	-
fryb	2666,3	806,4	2751,4	45,9	174,7	Endocrine	-
fstl4	609,4	309,8	1244,9	4,1	0,0	Endocrine	-
fut8a	1999,0	876,3	1688,7	22,2	231,3	Endocrine	-
fyna	1678,2	797,7	982,5	16,8	54,4	Endocrine	-
gabra2a	1166,4	6,0	328,2	1,5	0,0	Endocrine	-
gabrb1	935,4	1,9	1069,5	22,8	6,2	Endocrine	-
galnt16	2623,4	918,9	682,9	6,0	4,9	Endocrine	-
gch1	1795,6	1785,2	1621,8	211,3	9,1	Endocrine	-
gdf6a	5590,3	4531,6	6243,1	76,0	39,4	Endocrine	-
gdi1	5722,8	6560,2	6130,0	705,8	63,1	Endocrine	-
gfra3	9805,8	6,5	590,4	12,2	178,6	Endocrine	-
ghrhrb	24,2	434,8	1104,6	27,6	0,0	Endocrine	-

glrbb	1308,8	1131,5	1617,7	11,8	57,2	Endocrine	-
glsa	1002,8	880,3	385,5	15,6	76,0	Endocrine	-
gnai2b	5567,9	4985,1	1434,4	66,4	110,2	Endocrine	-
gnao1a	3149,2	3334,6	2133,9	29,2	67,1	Endocrine	-
gnao1b	3150,5	3865,4	1490,1	52,7	34,1	Endocrine	-
gnb5a	307,1	923,2	1156,9	5,4	6,5	Endocrine	-
gnb5b	1255,7	1667,3	525,4	4,5	0,0	Endocrine	-
gng13b	2071,7	2347,4	530,8	80,7	9,1	Endocrine	-
gpc1a	2297,1	7184,6	8131,9	12,3	48,5	Endocrine	-
gpc6a	250,2	1272,6	1082,7	0,5	139,2	Endocrine	-
gpd1b	10226,7	8143,9	4040,9	323,9	660,7	Endocrine	-
gpm6aa	4694,3	3693,1	2706,3	24,9	9,8	Endocrine	-
gpr137bb	1039,9	6414,3	780,0	57,3	148,6	Endocrine	-
gpr158a	1127,2	2312,1	31,9	1,6	0,0	Endocrine	-
gpr186	192,4	5,8	1128,2	0,3	0,0	Endocrine	-
gpr22a	4295,7	7566,7	686,9	45,3	139,5	Endocrine	-
gpt2l	17890,3	38313,2	2244,9	75,5	581,0	Endocrine	-
gpx3	61231,4	32432,4	9161,7	129,4	4692,8	Endocrine	-
gria1a	1791,7	282,2	39,6	21,5	49,9	Endocrine	-
gria2b	7115,1	3504,4	789,8	39,1	39,2	Endocrine	-
gria4a	1313,6	843,6	993,3	5,1	9,1	Endocrine	-
gria4b	1615,1	822,2	58,7	2,6	9,0	Endocrine	-
grid1a	1428,7	2351,0	684,2	10,1	10,7	Endocrine	-
grid1b	3690,3	1587,7	42,0	5,0	11,4	Endocrine	-
GRIK2	832,1	1155,9	3275,0	12,7	7,3	Endocrine	-
hhla2a.2	778,5	1697,2	937,5	9,8	28,4	Endocrine	-
hmgcra	475,3	2296,9	224,3	102,9	44,5	Endocrine	-
hprt1l	723,1	444,3	1020,6	12,8	9,7	Endocrine	-
hrh2b	6508,0	2871,5	151,2	18,9	76,7	Endocrine	-
hrh3	1206,0	1079,9	238,0	5,6	14,6	Endocrine	-
hs3st3b1b	5224,4	8253,5	4179,5	680,8	59,3	Endocrine	-
hsd17b12a	1474,1	1074,1	1063,4	45,2	4,1	Endocrine	-
hspa4l	1826,2	1366,7	1819,0	10,5	36,9	Endocrine	-
igfbp2b	2216,8	3410,0	1230,2	21,2	55,5	Endocrine	-
il11ra	4120,7	3628,1	3120,9	111,7	432,6	Endocrine	-
impdh1a	744,3	1095,3	907,6	1,9	3,2	Endocrine	-
inpp4b	235,3	603,0	1168,9	7,0	22,7	Endocrine	-
insm1a	4462,3	2355,9	9094,5	179,4	17,1	Endocrine	-
insm1b	1796,4	1563,9	1238,0	35,1	8,1	Endocrine	-
irf2bpl	6674,8	6147,7	5813,7	848,3	94,4	Endocrine	-
irs2a	4908,4	1227,9	3318,7	93,3	378,9	Endocrine	-
isl1	7091,0	12062,4	19012,2	59,3	67,4	Endocrine	-
itga1	431,2	44,3	1185,8	5,4	25,7	Endocrine	-
itga6a	1662,3	816,3	159,1	12,3	5,8	Endocrine	-
itm2ca	5526,4	4985,5	3224,3	16,4	9,0	Endocrine	-
itm2cb	2628,5	214,3	2554,8	160,8	59,4	Endocrine	-
jagn1a	486,8	1278,1	945,0	10,7	3,3	Endocrine	-
jam3b	2948,2	1587,8	1388,9	99,1	26,7	Endocrine	-
kcnc1a	2293,8	1040,8	1504,4	24,5	6,5	Endocrine	-
kcnc4	1652,8	757,1	853,8	1,3	11,4	Endocrine	-
kcnd3	301,9	499,2	1172,2	4,7	3,3	Endocrine	-
kcnh6a	1921,2	340,2	787,8	42,6	4,9	Endocrine	-
kcnip1b	791,1	5,0	1962,5	117,8	2,5	Endocrine	-

kcnip3b	2002,2	1193,5	523,3	71,9	66,6	Endocrine	-
kcnj11	454,2	2673,0	584,4	17,9	65,1	Endocrine	-
kcnj19a	2183,0	8494,2	410,8	7,9	55,3	Endocrine	-
kcnj19b	4388,1	3361,7	2521,6	152,8	26,9	Endocrine	-
kcnk9	1477,7	806,2	1164,3	3,6	2,5	Endocrine	-
kcnn3	1692,7	597,7	1335,2	10,0	78,1	Endocrine	-
kctd8	794,5	1055,1	858,4	14,2	7,5	Endocrine	-
kdm7ab	3873,1	4439,3	6092,3	579,9	307,9	Endocrine	-
kiaa1549la	1255,7	1697,9	1647,8	18,2	98,9	Endocrine	-
kif5bb	4135,2	5568,1	3765,5	151,0	634,8	Endocrine	-
kl	1137,3	16,3	8299,2	0,7	61,8	Endocrine	-
klf7a	842,7	1656,1	647,7	17,2	71,7	Endocrine	-
kremen1	1748,9	240,2	663,2	0,5	18,6	Endocrine	-
lepr	6664,2	3203,0	1699,6	245,1	219,8	Endocrine	-
lgi2a	4936,3	1654,2	193,1	0,2	127,0	Endocrine	-
lhfpl3	2021,0	863,1	1153,4	31,7	4,1	Endocrine	-
lin7a	1196,9	1345,0	1139,0	17,0	2,5	Endocrine	-
lmo1	1847,9	2318,7	1919,7	1,2	7,0	Endocrine	-
LO017739.1	665,9	903,4	1264,9	1,5	6,5	Endocrine	-
LO018188.1	409,6	374,2	1305,1	107,0	13,0	Endocrine	-
lrfn4a	2296,3	1923,3	1717,7	16,1	30,9	Endocrine	-
lrrc4ba	3709,1	1454,1	729,2	3,4	29,4	Endocrine	-
LRRC75A	1032,1	458,3	449,8	11,6	25,4	Endocrine	-
lrrn1	3469,4	824,6	4035,9	227,4	8,1	Endocrine	-
lrrn3b	308,5	91,8	1172,3	1,3	0,0	Endocrine	-
luzp2	1158,6	18,9	844,0	10,8	8,6	Endocrine	-
MANEAL	1138,4	391,1	630,2	10,5	6,5	Endocrine	-
map2	5365,5	3254,8	2798,4	20,3	34,5	Endocrine	-
map7d2b	2248,4	1507,6	1179,5	5,9	23,5	Endocrine	-
mapk4	2519,5	854,8	4101,5	21,3	56,3	Endocrine	-
mapk8ip3	4210,9	2007,3	1490,0	39,1	50,1	Endocrine	-
mast1b	1754,6	2099,5	539,1	8,8	6,5	Endocrine	-
mc5ra	129,9	1116,7	4188,0	0,9	4,9	Endocrine	-
meis1a	2211,4	3991,9	67,6	64,6	231,3	Endocrine	-
mical3a	1200,5	702,3	813,5	73,0	78,0	Endocrine	-
mllt11	356,4	3417,3	3701,9	18,3	13,6	Endocrine	-
mmp11b	1383,2	1287,9	227,7	5,5	0,0	Endocrine	-
mnx1	2595,1	4825,3	45,0	8,6	343,7	Endocrine	-
mtmr1a	1067,5	912,1	887,2	71,7	7,4	Endocrine	-
myo3b	1247,6	379,7	513,8	26,8	70,3	Endocrine	-
myo5aa	1896,4	1029,1	642,5	10,6	16,3	Endocrine	-
myt1b	1091,5	800,6	408,3	7,6	8,1	Endocrine	-
nalcn	3566,3	3307,4	3464,5	15,5	65,7	Endocrine	-
nav2b	3665,9	1827,6	2005,4	17,5	226,4	Endocrine	-
ncam1a	702,8	135,1	1004,7	4,6	54,2	Endocrine	-
ncam1b	998,8	1607,5	506,1	2,9	5,4	Endocrine	-
nckap5l	1153,9	571,4	403,1	44,1	0,0	Endocrine	-
ndrg4	1439,1	1904,2	1691,0	3,8	12,7	Endocrine	-
nell2a	252,2	3769,7	1705,3	0,8	0,0	Endocrine	-
neo1b	1009,5	624,3	426,3	30,6	4,1	Endocrine	-
neurod1	55043,8	31981,4	19794,5	215,9	373,9	Endocrine	-
nexmifb	1296,1	1972,4	659,1	51,1	53,2	Endocrine	-
nfasca	4272,1	3944,2	1234,3	213,5	81,2	Endocrine	-

nhsl2	789,5	242,4	1263,6	31,8	14,8	Endocrine	-
nmnat2	1768,3	2400,1	2833,2	13,3	19,7	Endocrine	-
nova2	812,2	384,7	1240,9	23,7	29,3	Endocrine	-
npas4a	8011,6	1182,0	10074,8	45,8	14,7	Endocrine	-
npffr1l3	25,3	1578,4	311,4	1,5	0,0	Endocrine	-
npr1a	4065,4	230,9	2561,8	167,6	20,3	Endocrine	-
nptx1l	1406,9	414,7	10,3	36,8	9,0	Endocrine	-
nrip2	3360,8	1131,5	75,6	74,7	67,4	Endocrine	-
nrsn1	4204,9	4235,7	3173,9	44,9	17,2	Endocrine	-
nrxn1a	6827,6	4259,7	5808,5	13,6	36,0	Endocrine	-
nrxn3a	1004,6	35,5	1252,3	10,4	7,5	Endocrine	-
nsfa	3101,0	2548,0	6593,2	10,3	29,4	Endocrine	-
nsg2	7334,1	4172,7	2358,6	111,6	55,0	Endocrine	-
ntrk3b	1219,5	2112,7	56,0	21,5	2,5	Endocrine	-
olfm1b	22292,1	16012,7	9466,5	69,7	124,0	Endocrine	-
oprd1b	890,4	3,8	4386,2	7,9	28,5	Endocrine	-
oprk1	2863,7	2338,3	1300,6	2,5	21,1	Endocrine	-
osbp2	1902,2	1679,5	158,6	12,9	0,0	Endocrine	-
osgin2	1595,0	502,9	1950,7	3,4	7,4	Endocrine	-
pacsin1a	1553,8	804,2	708,2	6,2	4,9	Endocrine	-
palm1a	80,8	427,9	1138,9	2,3	18,7	Endocrine	-
palm1b	621,5	1396,0	574,3	16,9	70,5	Endocrine	-
pax6b	18264,3	14624,1	30380,7	100,1	103,7	Endocrine	-
pcdh15b	1207,4	1429,1	48,9	17,0	6,5	Endocrine	-
pcdh20	1834,5	2119,2	656,6	49,8	5,4	Endocrine	-
pcloa	3531,3	1834,4	3582,3	42,4	17,6	Endocrine	-
pclob	1194,6	831,2	1403,5	23,2	16,1	Endocrine	-
pcnxl2	589,6	1009,3	263,0	73,1	12,7	Endocrine	-
pcp4a	852,6	1452,6	581,0	6,2	16,1	Endocrine	-
pcsk1	33776,7	89403,6	8955,9	26,2	350,4	Endocrine	-
pcsk1nl	26052,6	24221,9	22939,4	174,8	358,9	Endocrine	-
pcxb	38928,1	23444,4	869,8	5,3	82,4	Endocrine	-
pde4bb	1097,3	366,0	275,2	2,5	10,5	Endocrine	-
pde4cb	402,9	1586,0	150,4	12,2	42,6	Endocrine	-
pdyn	554,9	5,8	2342,4	5,8	6,6	Endocrine	-
pgr	867,7	22,1	1293,3	22,7	6,2	Endocrine	-
phgdh	1574,4	4085,1	1736,0	30,5	391,9	Endocrine	-
phyhipla	2158,6	2009,0	2057,8	12,6	8,1	Endocrine	-
pik3ip1	6720,1	10914,8	5095,4	685,2	93,8	Endocrine	-
pitpnab	2091,5	1566,2	980,1	4,9	22,7	Endocrine	-
plekhg5a	1367,9	1124,2	299,3	13,8	54,9	Endocrine	-
plppr5b	2014,0	3289,9	791,0	4,4	12,3	Endocrine	-
ppp3ca	827,5	1669,9	1985,4	6,0	31,1	Endocrine	-
prickle2a	356,8	357,1	1220,8	40,6	5,7	Endocrine	-
prkar1b	1712,1	1502,7	1877,6	12,4	53,2	Endocrine	-
prkcea	2803,2	2720,1	2781,0	219,0	118,6	Endocrine	-
prlra	5713,9	11470,6	7809,4	665,0	20,7	Endocrine	-
prokr1a	2313,7	1147,3	249,2	5,6	80,9	Endocrine	-
prune2	1260,5	708,5	672,6	63,8	0,0	Endocrine	-
ptger3	408,0	2308,3	3628,8	30,7	44,3	Endocrine	-
ptgs2a	125,1	37,9	8018,0	38,3	25,6	Endocrine	-
ptn	18354,2	12996,6	515,8	27,6	161,6	Endocrine	-
ptpn20	1744,9	1026,2	791,0	50,1	23,4	Endocrine	-

ptprdb	1992,5	1089,0	402,0	12,9	27,5	Endocrine	-
ptprn2	1506,4	1855,0	1955,4	12,6	10,9	Endocrine	-
ptprna	8502,4	3931,5	4498,2	131,4	928,5	Endocrine	-
ptprnb	9457,5	10428,1	3069,5	31,3	60,5	Endocrine	-
ptprua	812,9	2172,8	54,9	195,2	4,1	Endocrine	-
qpct	1652,3	1596,2	937,8	5,4	11,4	Endocrine	-
rab3c	78,2	2515,0	792,6	11,6	33,2	Endocrine	-
rab3da	1824,4	1397,1	3470,1	7,8	295,1	Endocrine	-
rab3db	1268,6	7,9	10041,2	194,1	63,7	Endocrine	-
rai1	1038,9	579,0	281,2	66,0	13,8	Endocrine	-
rapgef4	4024,4	2419,4	1712,2	27,0	13,9	Endocrine	-
rassf2a	1091,5	358,8	609,7	38,0	53,1	Endocrine	-
rcan2	3026,0	6758,0	1123,2	11,4	18,9	Endocrine	-
rcan3	1547,9	497,1	1990,8	20,0	18,0	Endocrine	-
renbp	3241,8	1017,0	448,7	8,3	187,7	Endocrine	-
ret	8341,0	4611,2	1997,9	89,6	26,1	Endocrine	-
rfx2	2916,7	1662,3	1613,8	299,1	65,8	Endocrine	-
rfx6	3409,6	2035,2	1723,2	9,2	37,1	Endocrine	-
rgs16	4817,8	8344,6	316,1	29,1	72,1	Endocrine	-
rgs3a	6615,1	5028,4	6100,7	86,4	139,1	Endocrine	-
rgs3b	1766,7	505,6	220,2	19,0	36,3	Endocrine	-
rgs4	10664,1	2783,8	197,1	453,1	91,0	Endocrine	-
rgs5a	132834,4	136957,5	18952,0	1469,9	1043,2	Endocrine	-
rgs8	1691,6	3078,1	414,8	13,4	24,3	Endocrine	-
rims2a	10474,8	16521,9	2795,9	83,4	61,4	Endocrine	-
rims4	1058,9	836,6	571,6	1,0	4,9	Endocrine	-
rnasekb	5194,6	13159,1	5830,6	444,1	953,5	Endocrine	-
rnf130	4375,0	2373,7	4883,8	214,2	518,2	Endocrine	-
rnf145a	2760,3	1534,0	1332,8	13,5	7,5	Endocrine	-
RNF157	1244,6	854,2	826,1	1,4	0,0	Endocrine	-
robo3	1550,2	1633,2	674,8	52,0	182,1	Endocrine	-
rprmb	2826,0	5898,6	3370,6	52,5	27,6	Endocrine	-
rps6ka1	4217,2	3378,7	3275,5	111,9	43,9	Endocrine	-
rtn1b	2487,2	8145,0	7888,0	30,9	108,6	Endocrine	-
rtn4rl2b	2389,2	17191,2	323,2	58,2	87,1	Endocrine	-
runx1t1	1172,2	275,4	576,6	6,9	6,5	Endocrine	-
rxfp3	1499,5	2202,1	445,8	16,4	9,8	Endocrine	-
scdb	3361,9	4139,9	4161,2	27,0	19,6	Endocrine	-
scg2b	8878,2	17782,6	3286,7	95,2	38,4	Endocrine	-
scg3	82967,3	153053,3	77077,8	1391,8	703,8	Endocrine	-
scg5	23909,0	30267,1	13266,8	218,0	148,2	Endocrine	-
scg5	2502,9	2352,2	1786,9	21,3	20,4	Endocrine	-
scgn	30225,3	64803,6	16607,7	545,9	384,4	Endocrine	-
scin	4078,0	910,9	2722,4	3,0	248,6	Endocrine	-
scinlb	24356,2	28,4	3515,6	265,3	441,4	Endocrine	-
scn3b	1363,7	891,2	570,5	5,4	4,9	Endocrine	-
sept5a	777,5	2189,1	632,3	16,4	13,8	Endocrine	-
sept9b	1733,9	2207,9	2384,9	14,4	227,8	Endocrine	-
serpina10a	4331,1	3969,6	2962,6	10,5	122,7	Endocrine	-
sesn3	18373,9	12501,1	14110,5	290,0	1654,1	Endocrine	-
sgpp1	1468,0	325,0	200,0	15,0	89,8	Endocrine	-
sgsm1a	3722,6	5127,6	1012,0	19,5	134,1	Endocrine	-
sgtb	1426,7	1244,3	1471,2	75,7	183,9	Endocrine	-

sh3bp5b	1482,4	316,2	504,4	6,0	10,6	Endocrine	-
sh3gl2a	9082,7	7024,3	3753,3	40,8	14,7	Endocrine	-
sh3glb2b	2152,8	2351,9	2577,3	168,0	11,2	Endocrine	-
shisa4	1869,8	1636,4	1275,6	148,2	41,9	Endocrine	-
shisal1a	1557,9	1692,5	639,6	19,1	58,3	Endocrine	-
si:ch211-106h4.9	2033,3	849,3	14,7	16,4	8,2	Endocrine	-
si:ch211-10a23.2	973,3	858,5	1217,7	10,6	31,5	Endocrine	-
si:ch211-113g11.6	7760,4	2844,6	2498,6	8,4	22,1	Endocrine	-
si:ch211-133n4.4	1825,0	2454,3	1869,0	10,3	2,5	Endocrine	-
si:ch211-137a8.4	1195,0	313,2	117,2	28,9	8,1	Endocrine	-
si:ch211-181d7.3	1235,6	18,4	708,4	5,2	7,4	Endocrine	-
si:ch211-195b13.1	10752,5	3433,8	1638,6	308,0	72,2	Endocrine	-
si:ch211-237l4.6	4843,3	7957,2	5495,9	92,4	161,6	Endocrine	-
si:ch211-242e8.1	1691,2	459,3	992,0	5,9	0,0	Endocrine	-
si:ch211-253p2.2	834,7	1201,6	52,7	83,6	0,0	Endocrine	-
si:ch211-255i3.4	2603,3	3121,1	971,8	121,3	140,9	Endocrine	-
si:ch73-119p20.1	1686,3	906,1	2133,1	1,5	7,4	Endocrine	-
si:ch73-287m6.1	2538,5	2456,2	1204,8	1,7	6,5	Endocrine	-
si:ch73-288o11.5	5885,2	1003,1	51,6	82,1	55,2	Endocrine	-
si:ch73-335m24.5	930,6	1624,8	1230,3	32,4	1,3	Endocrine	-
si:dkey-110c1.10	1855,0	1510,7	1094,5	202,9	14,1	Endocrine	-
si:dkey-110c1.7	2506,3	1149,2	1116,6	3,7	160,4	Endocrine	-
si:dkey-153k10.9	9504,7	13531,0	3976,7	143,9	314,2	Endocrine	-
si:dkey-183i3.9	1865,9	894,2	50,4	9,2	2,5	Endocrine	-
si:dkey-201i24.6	718,7	808,4	2091,8	17,5	17,5	Endocrine	-
si:dkey-238c7.16	119,3	804,9	1568,6	5,6	2,6	Endocrine	-
si:dkey-253d23.4	1128,7	1080,3	386,3	21,9	68,9	Endocrine	-
si:dkey-40m6.8	1105,1	672,1	614,3	22,7	13,4	Endocrine	-
si:dkey-81l17.6	226,1	639,6	1133,4	36,2	4,1	Endocrine	-
si:dkey-91m11.5	1303,4	248,3	1769,0	24,8	49,2	Endocrine	-
si:dkeyp-14d3.1	1328,1	2,3	448,9	0,7	0,0	Endocrine	-
sipa1l2	1257,7	907,6	482,8	47,5	37,3	Endocrine	-
slc12a7b	1173,0	495,7	1538,4	102,6	57,2	Endocrine	-
slc16a9b	5772,8	1685,8	1481,9	128,5	82,4	Endocrine	-
slc18a2	2691,8	1639,8	43,6	0,3	9,9	Endocrine	-
slc25a25b	1735,6	1516,3	315,7	102,0	132,3	Endocrine	-
slc25a6	2059,6	1388,9	1034,4	14,7	142,1	Endocrine	-
slc30a2	4437,5	8632,9	126,9	12,0	36,5	Endocrine	-
SLC35G1	2402,6	1748,9	835,3	194,8	37,5	Endocrine	-
slc35g2a	3392,4	2661,1	4891,7	29,5	47,5	Endocrine	-
slc43a2a	912,4	641,1	5722,4	23,0	40,8	Endocrine	-
slc43a2b	15413,2	13296,2	3231,1	420,1	101,8	Endocrine	-
slc6a11b	218,7	5,4	1474,0	0,4	9,7	Endocrine	-
slc6a17	484,2	1177,2	5129,3	17,2	55,3	Endocrine	-
slc7a4	910,7	1047,6	1782,3	15,1	10,1	Endocrine	-
slc9a7	1133,5	739,5	753,0	70,7	17,0	Endocrine	-
slco5a1a	3937,9	1508,6	952,9	6,7	25,1	Endocrine	-
slit2	1558,0	217,9	772,8	67,1	101,8	Endocrine	-
smarcd3a	1694,6	859,2	552,1	2,0	4,3	Endocrine	-
snap25a	10189,3	5956,1	9878,4	82,4	192,7	Endocrine	-
SNAP91 (1 of many)	2856,9	3137,2	2248,9	26,9	9,8	Endocrine	-
sncgb	2280,4	1903,5	100,3	1,2	7,4	Endocrine	-
snx10a	14,1	1820,0	2042,0	25,5	11,6	Endocrine	-

SORCS3	1074,3	516,2	18,6	2,3	5,0	Endocrine	-
spata22	1920,0	2471,6	1084,9	120,6	52,7	Endocrine	-
spon1b	782,4	22791,7	4990,1	414,2	191,3	Endocrine	-
sprn2	1737,8	665,5	154,1	4,0	2,5	Endocrine	-
spsb4a	731,6	562,4	1034,9	140,0	11,5	Endocrine	-
st6gal2a	1220,9	358,1	875,5	45,4	6,6	Endocrine	-
stat5a	1366,2	1725,5	1285,9	70,2	214,9	Endocrine	-
stk17a	918,9	780,4	1559,5	10,4	82,1	Endocrine	-
stmn2a	1532,7	1654,0	248,6	7,6	6,5	Endocrine	-
stx12l	1223,5	584,4	1161,9	6,5	2,5	Endocrine	-
stx1b	2755,7	2183,6	2317,0	20,6	12,0	Endocrine	-
stxbp1a	5298,2	3913,9	3863,0	13,4	68,1	Endocrine	-
stxbp5a	2738,0	1420,5	6874,3	36,4	125,9	Endocrine	-
stxbp5l	1122,3	320,2	735,7	3,7	7,4	Endocrine	-
sv2a	4440,9	2989,1	433,4	20,7	50,9	Endocrine	-
sv2bb	3702,4	3,0	1918,2	12,0	13,2	Endocrine	-
sybu	8363,3	367,3	2998,3	0,2	64,6	Endocrine	-
syngr3b	4664,9	4507,7	5964,9	49,4	59,1	Endocrine	-
sypa	411,4	1018,1	1041,1	65,7	2,5	Endocrine	-
syt11a	5203,7	4027,8	6897,0	46,0	19,6	Endocrine	-
syt14a	579,9	1054,8	742,6	2,9	0,0	Endocrine	-
syt1a	6541,3	6062,2	6579,9	136,3	83,9	Endocrine	-
syt4	12739,1	2642,2	14570,6	13,0	97,2	Endocrine	-
tanc2a	6578,7	8749,3	6988,8	324,1	87,2	Endocrine	-
TENM3	321,9	27,5	1070,9	32,2	2,5	Endocrine	-
tfeb	1194,6	1720,5	605,8	180,3	50,2	Endocrine	-
thsd7ba	1399,5	80,3	593,7	1,7	4,9	Endocrine	-
tle2a	543,5	259,6	1251,9	21,9	64,0	Endocrine	-
tmem106bb	1321,8	1037,8	757,7	65,6	103,3	Endocrine	-
tmem145	2269,4	750,7	379,7	22,1	20,1	Endocrine	-
tmem178b	629,2	1112,9	1054,7	26,9	96,9	Endocrine	-
tmem35	2325,6	4026,7	3225,6	42,0	16,6	Endocrine	-
tmem59l	2194,2	1800,5	2495,9	17,5	56,3	Endocrine	-
tmem63c	4332,7	1820,9	4638,4	581,2	7,4	Endocrine	-
tnc	2931,2	1189,9	10645,4	5,1	9,8	Endocrine	-
tnfaip2a	2244,6	1620,2	3272,2	48,4	20,5	Endocrine	-
tox	7194,8	4264,5	9093,9	39,0	26,1	Endocrine	-
trim2b	325,2	207,3	1644,1	31,0	87,6	Endocrine	-
trmt9b	2092,2	39,0	732,2	18,3	2,5	Endocrine	-
tshz2	704,4	635,8	1257,8	88,7	1,6	Endocrine	-
tspan7b	16184,7	12888,8	11789,7	121,8	111,1	Endocrine	-
ttll10	3266,5	2132,1	4390,2	309,5	138,5	Endocrine	-
ugcg	1518,9	1008,9	1024,6	59,3	123,7	Endocrine	-
ulk1b	1797,7	770,5	841,7	32,3	105,5	Endocrine	-
unc80	1016,0	948,9	956,1	4,7	7,5	Endocrine	-
ush1c	609,6	63,8	2066,5	3,2	0,0	Endocrine	-
uts1	32,9	3308,4	3533,7	9,6	0,0	Endocrine	-
vamp1	1642,9	3207,1	2629,9	246,1	12,3	Endocrine	-
vamp2	2460,9	1989,9	2769,3	19,1	12,8	Endocrine	-
vat1	5973,6	4971,3	7311,4	44,9	152,5	Endocrine	-
vat1l	262,6	1158,3	567,1	9,0	0,0	Endocrine	-
vcla	709,6	1536,6	151,8	11,3	16,2	Endocrine	-
vil1	2639,7	2996,0	33,3	91,4	17,8	Endocrine	-

vill	3828,7	774,2	0,3	13,6	6,5	Endocrine	-
vwa5b2	18012,7	9067,8	8773,0	29,1	136,4	Endocrine	-
wdfy3	1244,2	889,3	2307,6	34,5	195,0	Endocrine	-
XKR4	1924,4	716,7	1123,9	6,2	4,9	Endocrine	-
ywhag1	367,8	1281,0	511,7	9,4	123,8	Endocrine	-
zbtb16a	605,1	1000,8	2158,2	163,8	71,2	Endocrine	-
zcchc24	1172,9	295,9	203,5	19,8	13,3	Endocrine	-
zgc:101731	1014,5	1450,8	649,1	190,2	2,6	Endocrine	-
zgc:122979	2237,2	515,5	869,0	18,7	21,8	Endocrine	-
zgc:153031	1023,7	1248,1	1541,0	70,7	4,9	Endocrine	-
zgc:153704	1842,5	521,0	19,9	8,0	0,0	Endocrine	-
zgc:171566	936,5	2152,7	1424,5	13,2	8,5	Endocrine	-
zgc:194578	1207,7	658,4	437,1	119,3	12,2	Endocrine	-
zgc:194678	1454,7	2793,1	88,2	4,3	2,5	Endocrine	-
zgc:194990	540,5	4396,2	27,5	15,5	27,2	Endocrine	-
zgc:92658	2396,3	3,5	489,0	0,0	2,5	Endocrine	-
znrf1	942,2	1504,7	567,4	107,3	13,0	Endocrine	-

Supplemental table 2

Set of significative marker genes (pvalue adjusted <= 0.1) identified for all cell types based on zebrafish scRNA-seq dataset.

Zebrafish SC markers

Gene	Pct.Cluster	Pct.Others	Avg_LogFC	Pval_Adjusted	Cluster
gcga	1	0,992	3,9922223746	9.99988867182683e-321	Alpha
gcgb	0,995	0,98	3,9742613696	9.99988867182683e-321	Alpha
pnoca	0,995	0,806	3,0534211386	9.99988867182683e-321	Alpha
npb	0,985	0,189	2,4564321914	9.99988867182683e-321	Alpha
pcp4a	0,896	0,231	1,1966969441	9.99988867182683e-321	Alpha
rgcc	0,62	0,068	1,1632647947	9.99988867182683e-321	Alpha
cfap58	0,683	0,046	0,7541059995	9.99988867182683e-321	Alpha
ak5l	0,642	0,056	0,6053846592	9.99988867182683e-321	Alpha
scinlb	0,642	0,103	0,4966394895	5,7069477676218E-247	Alpha
slc7a2	0,819	0,208	0,6617015782	8,8597729150749E-235	Alpha
olfm1b	0,989	0,727	0,8660939351	1,5163030057432E-223	Alpha
rgs4	0,604	0,11	0,4772948987	2,0333582841976E-198	Alpha
krt15	0,791	0,259	0,8256715789	2,9108325815209E-185	Alpha
bckdk	0,837	0,335	0,7642668853	6,6955364353264E-181	Alpha
tdh	0,536	0,116	0,9419574269	2,2734725027731E-156	Alpha
fev	0,911	0,429	0,7335422041	1,9544854835083E-155	Alpha
cntnap2a	0,62	0,161	0,4068948778	9,9027964959245E-142	Alpha
pkib	0,51	0,108	0,3056064614	1,0801809105089E-138	Alpha
pcxb	0,521	0,134	0,3394822113	2,2232418133087E-114	Alpha
gpd1b	0,804	0,465	0,3484684325	8,38063064501721E-57	Alpha
slc43a2b	0,811	0,516	0,324188032	2,20355707856382E-40	Alpha
ssr3	0,871	0,662	0,3116076241	3,36828241783608E-40	Alpha
ins	1	0,998	3,0457515849	9.99988867182683e-321	Beta
ndufa4l2a	0,939	0,161	1,1661070909	9.99988867182683e-321	Beta
cxcl14	0,87	0,244	1,0010330525	9.99988867182683e-321	Beta
mnx1	0,59	0,091	0,5168111414	9.99988867182683e-321	Beta
krt8	0,821	0,375	0,4118310248	4,7446617919164E-227	Beta
herpud1	0,595	0,176	0,5831301573	7,7531270355501E-214	Beta
ctsla	0,819	0,429	0,5363587426	1,428575886763E-182	Beta
gstp1	0,581	0,186	0,4207569673	6,623216798577E-170	Beta
rims2b	0,518	0,15	0,4512205264	2,0873731324191E-159	Beta
tcima	0,663	0,272	0,3648225482	5,1900121217159E-157	Beta
zfand2a	0,647	0,264	0,4427341148	1,5902128664272E-142	Beta
tagln3b	0,796	0,477	0,4791217862	6,4134397476433E-139	Beta
hsp90b1	0,879	0,654	0,5730491216	3,2900829318446E-138	Beta
hspa5	0,97	0,878	0,5770665963	5,5363673399933E-136	Beta
tspan36	0,53	0,208	0,3252240741	6,9829036510921E-121	Beta
pim3	0,707	0,413	0,4658207084	9,4951094030096E-108	Beta
mknk2b	0,775	0,477	0,3641803261	4,19122690363861E-98	Beta
ctsd	0,63	0,341	0,4073548477	2,9246825093755E-95	Beta
tp53inp1	0,542	0,258	0,29972038	5,26338333841936E-86	Beta
syngr2b	0,614	0,324	0,2985894556	1,37142391962322E-81	Beta
zgc:153981	0,542	0,264	0,2904604035	6,05504214494373E-80	Beta
cebpd	0,63	0,34	0,3127547445	5,05354266800697E-77	Beta
si:ch211-260e23.9	0,562	0,316	0,3630309118	7,25445734223115E-69	Beta
uts1	0,616	0,36	0,4081044475	2,90007756724264E-65	Beta
fkbp4	0,764	0,483	0,3225404347	3,60582328985037E-61	Beta
calr	0,81	0,633	0,3098398402	1,24890582656971E-59	Beta
cbx7a	0,8	0,554	0,3894188718	7,67561813219104E-59	Beta
hspd1	0,742	0,499	0,2576995669	3,5726034924776E-58	Beta
sqstm1	0,64	0,424	0,3593469259	2,48524260191634E-55	Beta
manf	0,639	0,447	0,2860978785	1,58375049883526E-47	Beta

sst2	0,991	0,989	3,884100018	9.99988867182683e-321	Delta1.2
sst1.2	0,99	0,987	3,4429698023	9.99988867182683e-321	Delta1.2
zgc:195023	0,957	0,105	2,242272583	9.99988867182683e-321	Delta1.2
jun	0,986	0,68	2,1294358632	9.99988867182683e-321	Delta1.2
cldni	0,94	0,099	2,1009597854	9.99988867182683e-321	Delta1.2
nmba	0,937	0,345	2,0241965602	9.99988867182683e-321	Delta1.2
kctd12.2	0,968	0,138	2,0092972494	9.99988867182683e-321	Delta1.2
map2k6	0,968	0,274	1,9615129098	9.99988867182683e-321	Delta1.2
fabp6	0,929	0,091	1,8567017861	9.99988867182683e-321	Delta1.2
si:ch211-195b15.8	0,868	0,106	1,6281827914	9.99988867182683e-321	Delta1.2
pitpnaa	0,914	0,146	1,6052440566	9.99988867182683e-321	Delta1.2
serpinh1b	0,827	0,341	1,4187821777	9.99988867182683e-321	Delta1.2
ccl20b	0,795	0,067	1,2886570223	9.99988867182683e-321	Delta1.2
slc7a7	0,813	0,09	1,2278603428	9.99988867182683e-321	Delta1.2
mdkb	0,824	0,054	1,2191732422	9.99988867182683e-321	Delta1.2
ier2b	0,822	0,308	1,1196915724	9.99988867182683e-321	Delta1.2
dhrs1	0,841	0,425	1,0897615832	9.99988867182683e-321	Delta1.2
ifitm5	0,749	0,031	1,0890340406	9.99988867182683e-321	Delta1.2
slc3a2a	0,846	0,527	1,0883897748	9.99988867182683e-321	Delta1.2
carhsp1	0,827	0,209	1,0543883077	9.99988867182683e-321	Delta1.2
alcamb	0,9	0,542	1,0293041578	9.99988867182683e-321	Delta1.2
gulp1b	0,75	0,039	1,0286438235	9.99988867182683e-321	Delta1.2
hopx	0,761	0,124	1,00744725	9.99988867182683e-321	Delta1.2
mrap2a	0,708	0,039	0,8825116519	9.99988867182683e-321	Delta1.2
olfm2a	0,632	0,031	0,8371611041	9.99988867182683e-321	Delta1.2
marcksl1b	0,587	0,07	0,8015057669	9.99988867182683e-321	Delta1.2
oaz1a	0,975	0,927	0,7861686976	9.99988867182683e-321	Delta1.2
oaz2a	0,799	0,362	0,7857090967	9.99988867182683e-321	Delta1.2
sesn1	0,908	0,653	0,7508903891	9.99988867182683e-321	Delta1.2
slc26a3.2	0,624	0,039	0,7417819507	9.99988867182683e-321	Delta1.2
map1ab	0,635	0,115	0,7354795562	9.99988867182683e-321	Delta1.2
kl	0,632	0,057	0,7325813457	9.99988867182683e-321	Delta1.2
slmapb	0,569	0,039	0,7081592868	9.99988867182683e-321	Delta1.2
rab3db	0,682	0,143	0,7006097037	9.99988867182683e-321	Delta1.2
anxa11a	0,624	0,132	0,693570985	9.99988867182683e-321	Delta1.2
ppp1r1b	0,515	0,024	0,6466681054	9.99988867182683e-321	Delta1.2
spock3	0,581	0,098	0,583371343	2,5349875547627E-303	Delta1.2
stxbp6	0,64	0,173	0,6313687882	7,4220880877898E-293	Delta1.2
phkg2	0,514	0,101	0,5419895279	1,9083098864295E-253	Delta1.2
junbb	0,846	0,585	0,7404483874	2,4480210733261E-241	Delta1.2
ier2a	0,645	0,264	0,9620217812	2,4322929701645E-232	Delta1.2
amd1	0,585	0,173	0,5882352011	1,0513357170062E-229	Delta1.2
si:ch211-131k2.3	0,571	0,172	1,0509175208	3,8888879446922E-225	Delta1.2
ppp1caa	0,667	0,292	0,5921906121	1,1263927400837E-217	Delta1.2
nt5dc2	0,581	0,181	0,525225098	1,3095310515855E-212	Delta1.2
oaz1b	0,939	0,878	0,453757496	1,0484057373548E-211	Delta1.2
BX548077.1	0,652	0,235	0,5657354234	4,3843837093807E-210	Delta1.2
nr4a1	0,845	0,592	0,6050968329	3,7282846578082E-207	Delta1.2
egr2a	0,553	0,157	0,7133580196	5,8953367497539E-206	Delta1.2
ckbb	0,71	0,359	0,5844794594	5,9297026617625E-205	Delta1.2
tob1a	0,811	0,589	0,6032691549	1,3301396138971E-198	Delta1.2
rasgef1ba	0,565	0,194	0,5181886617	2,8617976088577E-195	Delta1.2
ndrg3b	0,552	0,195	0,4958341249	7,5328208984055E-192	Delta1.2

c2cd4a	0,955	0,91	0,516648926	8,5804997124468E-184	Delta1.2
mapk14a	0,532	0,18	0,442604873	1,0417275413136E-181	Delta1.2
syt4	0,655	0,35	0,5747066972	1,0934844272204E-175	Delta1.2
midn	0,598	0,276	0,5140169129	4,2298104001536E-162	Delta1.2
ppp6c	0,536	0,215	0,4581453692	5,5253476096817E-160	Delta1.2
fam107b	0,621	0,294	0,5185508089	1,4736026417449E-157	Delta1.2
ap3s2	0,691	0,427	0,5058688388	6,2181892308983E-156	Delta1.2
atf3	0,707	0,44	0,9394503623	1,0286809315307E-153	Delta1.2
ugp2b	0,519	0,215	0,4184550531	2,040516927324E-147	Delta1.2
fosb	0,603	0,292	0,5161366045	3,6553601129342E-142	Delta1.2
jund	0,595	0,29	0,4886736621	8,3990202091599E-142	Delta1.2
ubc	0,812	0,682	0,4849166656	1,8749767105906E-124	Delta1.2
srsf5a	0,855	0,752	0,4007805165	3,4711781886646E-121	Delta1.2
gabarapl2	0,784	0,615	0,3951574058	4,0875693958728E-116	Delta1.2
atpv0e2	0,68	0,461	0,3763132417	1,4631708826032E-109	Delta1.2
cd81a	0,537	0,272	0,3823933281	1,531171418262E-108	Delta1.2
uchl1	0,732	0,533	0,8219733077	1,0940035907321E-106	Delta1.2
hmgb1b	0,653	0,424	0,3955330224	1,1080143190755E-105	Delta1.2
ier5	0,547	0,298	0,418279689	9,2492212886545E-102	Delta1.2
mcl1a	0,988	0,967	0,4972126934	4,7943981951077E-101	Delta1.2
pax6b	0,794	0,64	0,378077821	1,81939150984608E-96	Delta1.2
sdcbp2	0,714	0,538	0,3533926635	4,16441228146524E-95	Delta1.2
hnrnpa0b	0,857	0,78	0,3893749005	1,17327312239255E-88	Delta1.2
fxyd1	0,977	0,969	0,2734805519	3,04409147273386E-86	Delta1.2
sat1a.2	0,783	0,691	0,5059847377	4,94018947893218E-86	Delta1.2
skp1	0,838	0,753	0,3171958055	1,17034451749601E-81	Delta1.2
pkma	0,597	0,395	0,3355623678	3,05473574686707E-81	Delta1.2
dynll1	0,94	0,907	0,3174744246	3,69070965609798E-81	Delta1.2
pim1	0,82	0,701	0,3810148458	4,53198373453945E-79	Delta1.2
dusp2	0,897	0,816	0,3066843611	1,61860031075438E-77	Delta1.2
rsrp1	0,972	0,949	0,3756942211	5,29686249851609E-75	Delta1.2
isl1	0,652	0,454	0,2920681246	6,50564719871874E-75	Delta1.2
zgc:77650	0,757	0,671	0,3842274221	7,82097305192622E-75	Delta1.2
rtn1b	0,647	0,495	0,3631985835	5,12700470555677E-74	Delta1.2
sik1	0,562	0,348	0,4001582039	1,02380742900437E-73	Delta1.2
mapkapk2a	0,566	0,361	0,3307679437	5,49842418697407E-73	Delta1.2
csnk1a1	0,66	0,511	0,3397982364	8,9848998763007E-72	Delta1.2
tspan7	0,713	0,579	0,3087215642	7,44677088318959E-71	Delta1.2
ndrg3a	0,644	0,485	0,3197898804	2,09993208553601E-70	Delta1.2
camk2n1a	0,875	0,786	0,3112179595	2,43279263050367E-70	Delta1.2
gstm.1	0,607	0,425	0,3434264897	6,08026175897209E-70	Delta1.2
calm2a.1	0,843	0,783	0,2873495823	2,75016840777609E-69	Delta1.2
cd164	0,673	0,479	0,3260009191	8,27064042843123E-69	Delta1.2
h3f3c	0,729	0,588	0,2730143493	1,19655889856968E-65	Delta1.2
vegfaa	0,517	0,315	0,2798905872	3,75279803310632E-64	Delta1.2
hist2h2l	0,908	0,847	0,3037062913	1,1889902417973E-62	Delta1.2
aplp2	0,507	0,313	0,3018592026	1,27626212026922E-62	Delta1.2
laptm4a	0,733	0,622	0,3111452754	1,7679583233776E-62	Delta1.2
pcbp2	0,593	0,424	0,2897448763	6,39565387575862E-60	Delta1.2
ip6k2a	0,515	0,329	0,2949809892	6,41859562664971E-60	Delta1.2
cope	0,749	0,663	0,2849060401	7,90695492248558E-60	Delta1.2
itm2ba	0,907	0,877	0,2546386333	6,15449684041699E-59	Delta1.2
arl6ip1	0,729	0,624	0,2875897815	2,45128178376801E-58	Delta1.2

selenot2	0,518	0,34	0,2823135635	5,39981286094768E-58	Delta1.2
ufc1	0,617	0,465	0,2737575134	5,71389238114984E-54	Delta1.2
laptm4b	0,706	0,628	0,5265036432	3,79207196794046E-53	Delta1.2
si:dkey-4p15.3	0,742	0,589	0,5592280276	5,63636865285829E-48	Delta1.2
mcl1b	0,748	0,634	0,2980754592	3,68295037415755E-47	Delta1.2
oser1	0,589	0,458	0,3065560571	1,72026288511628E-44	Delta1.2
hsp70.2	0,992	0,989	0,3063200658	5,97538504441614E-43	Delta1.2
cdc42	0,57	0,464	0,2573007839	9,25745483356273E-35	Delta1.2
hspa4a	0,738	0,672	0,3455517382	9,22602346083941E-32	Delta1.2
hsp70.1	0,915	0,963	0,7236476333	2,6506780776101E-24	Delta1.2
sst1.1	0,999	0,992	4,186508731	9.99988867182683e-321	Delta1.1
sox11b	0,723	0,09	1,2762714636	9.99988867182683e-321	Delta1.1
srd5a2b	0,909	0,378	1,2853012252	1,8544905700256E-273	Delta1.1
nme2b.1	0,993	0,98	0,6668912929	6,6663339786322E-229	Delta1.1
hsp90ab1	0,999	0,992	0,6189268929	1,6734783284758E-218	Delta1.1
scgn	0,98	0,904	0,9651982297	3,8133192038658E-214	Delta1.1
faua	0,997	0,988	0,587356555	7,9891703966758E-201	Delta1.1
RPL41	0,999	0,999	0,6407766916	1,0095284333163E-198	Delta1.1
uba52	0,977	0,951	0,6647058601	1,1484229127645E-196	Delta1.1
zgc:114188	0,989	0,98	0,6222493354	6,1613185908957E-196	Delta1.1
eef1g	0,992	0,963	0,6102363184	2,8721374236198E-185	Delta1.1
btf3	0,962	0,913	0,6904776868	9,4536287837855E-174	Delta1.1
tpt1	1	0,997	0,4416527455	3,0916017423532E-163	Delta1.1
ppiaa	0,979	0,959	0,5696166628	7,0298565156052E-162	Delta1.1
meis1b	0,524	0,13	0,5867672003	7,4654834928525E-162	Delta1.1
eef1a1l1	1	0,999	0,4904196956	1,1982804700987E-158	Delta1.1
eef1b2	0,993	0,984	0,4892854922	3,254196669307E-157	Delta1.1
scg3	0,996	0,976	0,6399573983	2,3652905395389E-154	Delta1.1
mtbl	0,698	0,309	0,8176060731	1,6599803565462E-153	Delta1.1
rsl24d1	0,957	0,896	0,6425284896	1,3697965944893E-151	Delta1.1
atp5mc3b	0,956	0,867	0,6302341643	1,4716209311961E-150	Delta1.1
aldocb	0,929	0,813	0,7201929558	2,3213884348176E-148	Delta1.1
ptmab	0,936	0,77	0,7557448116	8,1185754009283E-145	Delta1.1
eif3f	0,969	0,917	0,6056674193	6,4528927289105E-142	Delta1.1
mt-atp6	0,992	0,981	0,6252546551	3,4583839941209E-138	Delta1.1
b2ml	0,945	0,828	0,6599143793	1,726757595325E-136	Delta1.1
tmsb4x	0,976	0,909	0,5767598606	1,4584385905767E-135	Delta1.1
rack1	0,993	0,986	0,471074484	2,4536269376443E-131	Delta1.1
pcsk1nl	0,96	0,925	0,6827020606	6,9420427816288E-128	Delta1.1
ppiab	0,996	0,987	0,4084906435	3,3364631204259E-119	Delta1.1
eef1db	0,936	0,862	0,6007588612	3,6361532388609E-117	Delta1.1
ndufa4l	0,883	0,711	0,618617543	4,042364057748E-117	Delta1.1
fabp3	0,799	0,549	0,6346931273	2,6270465693315E-107	Delta1.1
cfl1	0,964	0,93	0,4841951158	7,2093879012358E-107	Delta1.1
mibp2	0,885	0,726	0,6010971616	9,7870016964269E-107	Delta1.1
gck	0,84	0,619	0,7560787329	1,2300615101371E-103	Delta1.1
rtn4rl2b	0,574	0,228	0,7322824574	1,5420437359228E-103	Delta1.1
cox6b2	0,938	0,86	0,5270498659	3,9116838489364E-103	Delta1.1
sec61g	0,989	0,991	0,3947582506	9,6572102077875E-102	Delta1.1
g6pcb	0,543	0,211	1,0006157712	5,1948830165968E-101	Delta1.1
rims2a	0,772	0,514	0,7796868892	6,9027724605319E-100	Delta1.1
SLC5A10	0,515	0,193	0,5626128144	1,19671745138187E-96	Delta1.1
rrad	0,529	0,199	0,7147978045	6,80672746144623E-96	Delta1.1

gtpbp4	0,83	0,668	0,6003109547	4,89728317163348E-94	Delta1.1
ybx1	0,966	0,927	0,4946906104	8,19407386333051E-91	Delta1.1
cox4i1	0,894	0,847	0,4592027955	4,13084279105441E-79	Delta1.1
mt-nd5	0,735	0,51	0,5772171015	2,78101546572265E-78	Delta1.1
pfkfb4b	0,555	0,274	0,6551644894	7,01499488420036E-77	Delta1.1
nsa2	0,744	0,522	0,5320870238	4,94752782813291E-76	Delta1.1
mgst3b	0,917	0,862	0,5931175015	9,55658306656575E-70	Delta1.1
ppt1	0,708	0,47	0,5278149387	9,98424756086194E-69	Delta1.1
eif3i	0,804	0,658	0,4690499235	2,49769304575603E-65	Delta1.1
hsp70l	1	0,999	0,4012444267	1,38890533063583E-63	Delta1.1
cpe	0,765	0,639	0,5445207476	2,7181161817986E-62	Delta1.1
hspa8	0,981	0,972	0,3505474812	4,33767115655552E-62	Delta1.1
serbp1a	0,84	0,729	0,4442943566	9,8481413964913E-60	Delta1.1
atp5l	0,869	0,817	0,4207426514	4,51813081937697E-59	Delta1.1
xbp1	0,928	0,903	0,5005595886	3,87523191612703E-58	Delta1.1
zgc:175264	0,515	0,265	0,4131500358	5,76327835288916E-58	Delta1.1
cox7b	0,869	0,82	0,3988117148	1,30228606657936E-57	Delta1.1
emc10	0,727	0,56	0,5002147402	3,42926593850052E-56	Delta1.1
eif4a1b	0,818	0,678	0,439152072	7,34439273176059E-56	Delta1.1
hmgn6	0,902	0,841	0,4151905366	5,02435530599691E-53	Delta1.1
ola1	0,708	0,53	0,4612235959	5,1930643027313E-52	Delta1.1
fth1a	0,942	0,925	0,3575825369	2,50047793237434E-51	Delta1.1
eif3m	0,826	0,735	0,4347521206	3,61472670302496E-51	Delta1.1
zgc:56493	0,925	0,912	0,3266874	4,79791944144688E-51	Delta1.1
cited4a	0,855	0,773	0,4896200209	7,16606312828823E-51	Delta1.1
cox5aa	0,827	0,769	0,4120887246	4,1072144494011E-50	Delta1.1
tpi1b	0,78	0,657	0,4125643366	8,47666335237439E-50	Delta1.1
COX5B	0,851	0,779	0,3880314448	9,73764298844409E-50	Delta1.1
alcama	0,688	0,526	0,4222564321	8,30164480446044E-46	Delta1.1
pno1	0,743	0,618	0,4043098363	1,89525879719582E-45	Delta1.1
epcam	0,59	0,347	0,5383528642	5,20356360454771E-44	Delta1.1
nop53	0,634	0,45	0,4320316713	5,29848808272669E-44	Delta1.1
ier3ip1	0,784	0,713	0,4228622484	7,67277482490181E-43	Delta1.1
atp5pd	0,898	0,88	0,3355250678	2,32466767488508E-42	Delta1.1
actb1	0,89	0,811	0,590061228	5,58508200146979E-42	Delta1.1
cd63	0,847	0,809	0,3932408732	3,31659186641468E-41	Delta1.1
eif5a	0,891	0,866	0,3466948794	9,82547620313025E-41	Delta1.1
ube2al	0,55	0,357	0,4026784968	3,71311785039036E-40	Delta1.1
adh5	0,532	0,334	0,3583156989	7,55977221800288E-40	Delta1.1
nap1l1	0,615	0,442	0,41236586	8,76992644969263E-40	Delta1.1
cox7c	0,818	0,737	0,3444033248	2,21066558788012E-39	Delta1.1
cited4b	0,751	0,636	0,628152354	3,67582060197941E-39	Delta1.1
eif3g	0,709	0,573	0,3890784213	2,80776788545778E-38	Delta1.1
atp5f1d	0,791	0,72	0,3663405639	1,62362550946771E-37	Delta1.1
pcbd1	0,783	0,712	0,3502177931	1,64066844202639E-36	Delta1.1
atp5pb	0,772	0,686	0,3796678296	1,04820453083691E-35	Delta1.1
lygl1	0,515	0,328	0,4141780592	9,37696571931914E-35	Delta1.1
ubb	0,95	0,949	0,3465822084	5,12334193179055E-34	Delta1.1
pnrc2	0,973	0,975	0,251398258	1,07720118692973E-33	Delta1.1
eif3ba	0,631	0,485	0,3913733724	2,68434828030993E-33	Delta1.1
cirbpb	0,92	0,905	0,3191796542	2,29420690072292E-32	Delta1.1
eif3ha	0,609	0,449	0,3499689036	3,34762208396659E-32	Delta1.1
eif5a2	0,74	0,628	0,3709552753	5,22089863080853E-32	Delta1.1

fkbp5	0,698	0,594	0,4165436044	1,09837003958107E-30	Delta1.1
ndufs4	0,529	0,364	0,3448233207	3,31856744125821E-30	Delta1.1
tmed7	0,712	0,636	0,4031240564	9,8671896136901E-30	Delta1.1
atp5pf	0,786	0,716	0,3376497835	4,1057014444937E-29	Delta1.1
cox6c	0,765	0,705	0,3312366195	2,07297684582399E-27	Delta1.1
SMIM30	0,643	0,524	0,3483336983	2,52013686639179E-27	Delta1.1
npc2	0,767	0,696	0,3494736718	3,31568454009388E-27	Delta1.1
pam	0,507	0,35	0,3147524885	2,86960052076289E-25	Delta1.1
gabarapb	0,855	0,846	0,2784421262	3,78212125825896E-25	Delta1.1
higd1a	0,839	0,833	0,3188180855	3,78743567790449E-25	Delta1.1
CABZ01079011.1	0,615	0,469	0,3105606931	6,52665625761616E-24	Delta1.1
cox6a1	0,806	0,776	0,2879356112	1,94625588959139E-23	Delta1.1
ppib	0,814	0,787	0,3065060231	2,96140059517276E-23	Delta1.1
eif2s3	0,625	0,513	0,3401246978	3,893790398148E-23	Delta1.1
DNAJA4	0,858	0,823	0,3617605974	7,1120414544018E-23	Delta1.1
si:ch211-222l21.1	0,863	0,838	0,2742552246	7,49850100979604E-23	Delta1.1
si:ch211-202a12.4	0,96	0,966	0,300127994	2,16598145045166E-21	Delta1.1
capns1b	0,823	0,807	0,2964594601	4,19481653404272E-21	Delta1.1
pitpnbl	0,583	0,47	0,3406401746	7,18902330091054E-21	Delta1.1
eif3d	0,588	0,471	0,3222759812	2,69020594924282E-20	Delta1.1
cox7a2a	0,674	0,609	0,3096231078	9,69120814802723E-20	Delta1.1
hsp70.3	0,991	0,992	0,2592799528	1,20086133167912E-19	Delta1.1
atp5meb	0,866	0,852	0,2580648776	1,23833408676916E-19	Delta1.1
atp5f1c	0,727	0,671	0,27548721	2,7616983516582E-19	Delta1.1
zgc:158852	0,668	0,576	0,3139762095	1,09513965363577E-18	Delta1.1
atp1b1a	0,796	0,759	0,2751352555	1,48785431305161E-18	Delta1.1
cox5ab	0,736	0,699	0,2850146436	2,9088799507855E-18	Delta1.1
mid1ip1b	0,606	0,483	0,4724534578	5,17457666303466E-18	Delta1.1
hmgn2	0,627	0,545	0,285786604	2,60511376494984E-17	Delta1.1
spcs3	0,673	0,618	0,2780459103	1,06952877358014E-16	Delta1.1
eif2s1b	0,525	0,416	0,2855359423	1,12140354914218E-16	Delta1.1
chchd2	0,747	0,705	0,2700760027	1,81121760150586E-16	Delta1.1
tomm20b	0,554	0,439	0,2640923441	1,87826975596512E-16	Delta1.1
pfdn5	0,548	0,435	0,2717067991	4,2481855986793E-16	Delta1.1
atf4a	0,546	0,427	0,2773417844	4,76525844743586E-16	Delta1.1
fam46c	0,558	0,448	0,3587027685	5,02139820492712E-16	Delta1.1
slc25a3b	0,702	0,639	0,2876295791	8,13168723541345E-16	Delta1.1
atp5f1e	0,747	0,689	0,2671578463	1,09703074828693E-15	Delta1.1
pabpc1a	0,783	0,751	0,2770096791	1,45352415427817E-15	Delta1.1
bzw1b	0,617	0,51	0,3585626677	9,14431992939935E-15	Delta1.1
atox1	0,668	0,609	0,2811898087	2,26741393788252E-14	Delta1.1
edf1	0,783	0,741	0,2504939476	2,65831707979864E-14	Delta1.1
ucp2	0,771	0,718	0,3594449775	2,7587981490473E-14	Delta1.1
ifrd1	0,751	0,706	0,3499966353	3,21517598738126E-14	Delta1.1
dnajb1b	0,981	0,987	0,2531123134	4,72117001530474E-14	Delta1.1
ndufa1	0,609	0,53	0,2842262284	5,69085139085367E-14	Delta1.1
slc1a4	0,692	0,631	0,2727282133	6,26238263709773E-14	Delta1.1
mdh2	0,747	0,712	0,2758623311	6,70941288501487E-14	Delta1.1
ndufb8	0,664	0,605	0,2529634081	6,73730073782619E-14	Delta1.1
eif3ea	0,655	0,609	0,2921727943	1,24054682654065E-13	Delta1.1
cct7	0,654	0,609	0,2776202718	1,25364876340145E-13	Delta1.1
hmgb2a	0,727	0,645	0,2745686193	1,32041039668302E-13	Delta1.1
eif4a1a	0,529	0,421	0,2651685483	2,4851989988212E-13	Delta1.1

fkbp2	0,646	0,577	0,2515535033	3,10132741488388E-13	Delta1.1
vdac2	0,626	0,549	0,2667595919	5,23039295044665E-13	Delta1.1
jpt1b	0,52	0,425	0,2808935579	1,00809816288888E-12	Delta1.1
vdac3	0,765	0,731	0,2733331403	3,46278905911303E-12	Delta1.1
atp5po	0,704	0,672	0,2618667924	4,55708078720781E-12	Delta1.1
khdrbs1a	0,807	0,797	0,2690077658	4,79935820424846E-12	Delta1.1
eif1axb	0,665	0,615	0,2506688391	6,30221000871283E-12	Delta1.1
cct6a	0,536	0,444	0,2698799437	7,53430172598287E-12	Delta1.1
h2afx1	0,95	0,964	0,283487183	1,86206308447378E-11	Delta1.1
zgc:162944	0,696	0,644	0,265882131	4,49704670565151E-11	Delta1.1
ell2	0,646	0,607	0,2522304863	8,18086761598099E-11	Delta1.1
sh3gl2a	0,579	0,515	0,2788895424	1,18290074125351E-10	Delta1.1
ndufb10	0,614	0,558	0,2510625115	3,55131700447821E-09	Delta1.1
atp6ap2	0,704	0,698	0,2892024021	4,17460596834873E-09	Delta1.1
zgc:158463	0,684	0,653	0,2533343329	4,7958661796307E-09	Delta1.1
fam46ba	0,59	0,549	0,3176164035	0,000000019	Delta1.1
myl12.1	0,682	0,669	0,2719730851	6,35567367219967E-08	Delta1.1
adm2a	0,611	0,605	0,2920914845	0,0001429431	Delta1.1
hspb1	0,605	0,573	0,4674046328	0,0009154574	Delta1.1
mboat4	0,758	0,004	1,7376400207	9.99988867182683e-321	Epsilon
plcxd3	0,703	0,043	1,131205266	3,7845011753036E-174	Epsilon
acsl1a	0,56	0,048	1,0312424754	8,81541116409968E-101	Epsilon
mlnl	1	0,229	5,7352373604	1,3794428137748E-100	Epsilon
brd1b	0,604	0,058	1,0408315523	6,49114664404083E-99	Epsilon
abhd17c	0,637	0,077	1,0012490949	1,38805892744594E-85	Epsilon
si:rp71-17i16.6	0,857	0,219	1,4292022326	6,17471856440673E-66	Epsilon
ghrl	1	0,966	4,6585683428	7,25439536830945E-56	Epsilon
ndrg2	0,692	0,191	0,9586179315	3,34269178295137E-41	Epsilon
asah1b	0,67	0,212	1,1806034446	1,95369435363391E-34	Epsilon
zgc:92818	0,901	0,714	1,0467636982	1,64978506142297E-24	Epsilon
ca6	0,516	0,147	0,6384668617	3,84573484880463E-22	Epsilon
plp1b	0,56	0,216	0,5938315476	2,23457348932627E-15	Epsilon
mt-cyb	0,989	0,99	0,4359051771	3,777167376015E-12	Epsilon
spry4	0,527	0,258	0,5540572431	1,3832960919365E-08	Epsilon
snap25a	0,725	0,606	0,6768997695	4,18306135918936E-07	Epsilon
gpx4a	0,626	0,49	0,661245991	0,0001234105	Epsilon
neurod1	0,659	0,564	0,5812035391	0,016414716	Epsilon
dynlrb1	0,681	0,623	0,3474299851	0,0706567628	Epsilon
ndufb9	0,56	0,434	0,4051691951	0,0915816999	Epsilon
dkk3b	0,828	0,214	2,1005364681	9.99988867182683e-321	Endocrine
rgs5a	0,803	0,158	1,9632503717	9.99988867182683e-321	Endocrine
egr4	0,765	0,098	1,7476977874	9.99988867182683e-321	Endocrine
ppdpfb	0,963	0,685	1,714522831	9.99988867182683e-321	Endocrine
scg5	0,923	0,088	1,5166076369	9.99988867182683e-321	Endocrine
nucb2b	0,888	0,203	1,429064657	9.99988867182683e-321	Endocrine
clu	0,894	0,272	1,2296846761	9.99988867182683e-321	Endocrine
chga	0,665	0,037	1,0626069643	9.99988867182683e-321	Endocrine
CABZ01073265.1	0,748	0,088	0,9932767295	9.99988867182683e-321	Endocrine
rnasekb	0,813	0,059	0,9869206608	9.99988867182683e-321	Endocrine
eno1a	0,767	0,069	0,9176268982	9.99988867182683e-321	Endocrine
gapdhs	0,975	0,551	0,7895745305	9.99988867182683e-321	Endocrine
lgals2a	0,702	0,094	0,7387337446	9.99988867182683e-321	Endocrine
calm1b	0,741	0,11	0,7265749688	9.99988867182683e-321	Endocrine

vamp2	0,625	0,023	0,6937383357	2,0501620149612E-299	Endocrine
atp6v0cb	0,679	0,102	0,6669925994	3,4289587121972E-287	Endocrine
vat1	0,629	0,051	0,6204011486	9,5844938535978E-272	Endocrine
arl4ab	0,72	0,151	0,7600352444	7,0459935720932E-257	Endocrine
scg2b	0,584	0,039	0,7634752431	7,2329382345719E-252	Endocrine
cplx2l	0,596	0,046	0,6283599886	5,3667090261522E-250	Endocrine
vamp1	0,569	0,03	0,5787011872	1,066570670819E-247	Endocrine
atp6v1g1	0,753	0,23	0,6229517973	1,5319374892054E-244	Endocrine
tox	0,548	0,025	0,5797945354	5,6137354675393E-241	Endocrine
si:ch211-237l4.6	0,635	0,084	0,5611531194	2,5962505425054E-233	Endocrine
lysmd2	0,602	0,07	0,5614014803	3,2461501747633E-233	Endocrine
pbxip1b	0,71	0,193	0,7070399055	9,4707730750032E-229	Endocrine
syt1a	0,533	0,026	0,5347010479	3,4413580814582E-227	Endocrine
tspan7b	0,527	0,024	0,5328503983	1,0845734796185E-226	Endocrine
nsg2	0,53	0,03	0,7031034632	7,4487474417936E-224	Endocrine
sec11a	0,802	0,327	0,5658974951	6,2001477127917E-223	Endocrine
pdyn	0,516	0,036	1,4865746117	5,399173535341E-216	Endocrine
calm3a	0,901	0,419	0,6409466843	6,8412451297327E-213	Endocrine
pik3ip1	0,594	0,084	0,5699974543	1,6042527994828E-212	Endocrine
ptn	0,512	0,03	0,8423084891	5,6235885794934E-209	Endocrine
rnaseka	0,64	0,128	0,4988021993	7,2017743616944E-208	Endocrine
reep5	0,773	0,31	0,5522741221	7,1465373902567E-204	Endocrine
spcs2	0,707	0,237	0,5300037049	7,2167488908468E-198	Endocrine
rpl22l1	0,961	0,788	0,5320513697	4,1340248292849E-197	Endocrine
eef1a1a	0,478	0,024	0,4508272599	9,2151311778287E-193	Endocrine
si:dkey-13i19.8	0,806	0,321	0,5143461299	1,8914196243825E-189	Endocrine
ap1s2	0,464	0,022	0,45178185	5,5703159438825E-188	Endocrine
rab2a	0,772	0,291	0,5211128031	1,9353791775235E-187	Endocrine
tmem35	0,453	0,015	0,4914196459	2,39516942545E-187	Endocrine
atp6v1e1b	0,597	0,116	0,4696840469	1,8457123644732E-185	Endocrine
bhlhe41	0,499	0,048	0,4759953277	5,1125852048772E-185	Endocrine
stmn1b	0,46	0,027	0,7806169406	3,506775135016E-183	Endocrine
eef2l2	0,486	0,049	0,4652131873	6,0125005968128E-183	Endocrine
tusc3	0,511	0,06	0,4773694461	1,5691637775618E-182	Endocrine
sdf4	0,599	0,146	0,4991495091	3,0682700805698E-178	Endocrine
atcaya	0,43	0,011	0,4262579722	7,5265422615035E-177	Endocrine
si:ch211-133n4.4	0,434	0,015	0,4328120362	2,5471193740014E-175	Endocrine
zgc:92606	0,58	0,128	0,4710247556	5,1245787468012E-172	Endocrine
serpina10a	0,443	0,026	0,4331834319	7,4914953095212E-172	Endocrine
calm1a	0,881	0,43	0,5062775964	4,2211927781007E-166	Endocrine
tsc22d1	0,533	0,098	0,4359135511	2,9187967252053E-161	Endocrine
fam49bb	0,431	0,029	0,4119125969	4,5206685210682E-160	Endocrine
ufm1	0,592	0,165	0,4283917398	2,1125979146407E-159	Endocrine
rprmb	0,395	0,011	0,4294351514	1,6734142003389E-157	Endocrine
elocb	0,707	0,262	0,4207021367	8,0982349628881E-157	Endocrine
zgc:65894	0,403	0,016	0,4280410381	8,5730007024381E-156	Endocrine
hint1	0,746	0,34	0,4371512197	2,6694777167117E-155	Endocrine
tuba2	0,417	0,027	0,4395515616	5,6724560545953E-155	Endocrine
ctsf	0,536	0,117	0,4303512569	5,8863770606867E-154	Endocrine
rgs16	0,422	0,032	0,6261598299	7,6575081623306E-153	Endocrine
dnajb4	0,566	0,138	0,483966886	3,7322479972004E-149	Endocrine
spon1b	0,42	0,037	0,4591239965	9,6066748854772E-149	Endocrine
bnip3lb	0,505	0,094	0,3854894943	6,8071506865761E-148	Endocrine

epdr1	0,386	0,015	0,3679751322	3,5365388689612E-147	Endocrine
prdx5	0,617	0,169	0,3960755165	5,5877269062528E-147	Endocrine
elovl1a	0,529	0,105	0,461275359	8,1182688874892E-147	Endocrine
atp5mc3a	0,699	0,295	0,4126539434	2,1502434128875E-145	Endocrine
cst3	0,944	0,654	0,7063579904	2,0147820792449E-144	Endocrine
tspan3a	0,685	0,237	0,4254198168	2,1773510182647E-144	Endocrine
ywhag2	0,381	0,021	0,3656870009	2,1713419712971E-140	Endocrine
gtf2h5	0,524	0,123	0,3873068682	1,0134662903019E-136	Endocrine
tuba1c	0,365	0,021	0,3881340522	1,1937639039844E-131	Endocrine
tspan13b	0,387	0,039	0,3407410088	1,2807685730518E-129	Endocrine
sec62	0,512	0,125	0,3720236325	2,3903618630064E-129	Endocrine
gpt2l	0,416	0,065	0,4915404627	3,7722025458953E-129	Endocrine
selenof	0,778	0,378	0,3712551856	3,8231789077235E-128	Endocrine
ap2s1	0,457	0,083	0,3420600873	6,8865907324381E-128	Endocrine
pfn2l	0,707	0,324	0,377916273	7,5077439348402E-128	Endocrine
tmbim4	0,607	0,2	0,3527405106	2,1237876273621E-127	Endocrine
fundc1	0,475	0,094	0,332994369	1,3540144573714E-126	Endocrine
ostc	0,606	0,226	0,3561691841	7,883813655303E-126	Endocrine
syngr3b	0,355	0,02	0,3401609539	1,6824523546767E-125	Endocrine
pdia6	0,496	0,124	0,379663889	2,1603506381035E-125	Endocrine
pyyb	0,375	0,039	1,1249678389	3,5108810082449E-125	Endocrine
apof	0,35	0,021	0,3739687312	5,6618452827563E-124	Endocrine
syt11a	0,332	0,013	0,3338640171	4,4063689841769E-122	Endocrine
ube2na	0,585	0,188	0,3104130329	3,0017234828289E-121	Endocrine
uqcc3	0,558	0,186	0,3542232881	2,2207349162749E-120	Endocrine
gng13b	0,331	0,015	0,3362829185	4,11671933572548E-119	Endocrine
pde6ha	0,33	0,015	0,7095437793	9,11446108859159E-119	Endocrine
slc35g2a	0,331	0,015	0,3413975217	5,2857947441367E-118	Endocrine
sub1a	0,755	0,341	0,3621281163	9,904786433669E-118	Endocrine
si:dkey-33c12.4	0,585	0,203	0,3464051313	1,3594004262768E-116	Endocrine
serinc1	0,478	0,118	0,3346779821	1,3631820352583E-116	Endocrine
si:dkey-81l17.6	0,327	0,015	0,3052620759	1,7789993944995E-116	Endocrine
si:ch211-199o1.2	0,346	0,027	0,4948988209	1,4305384249946E-115	Endocrine
tmed9	0,513	0,151	0,3332241569	6,0541587323305E-115	Endocrine
calr3b	0,461	0,102	0,331791293	9,6404215441221E-115	Endocrine
ube2v2	0,757	0,362	0,3367070981	2,0980970853376E-114	Endocrine
si:dkeyp-75h12.5	0,36	0,04	0,3665407536	7,4359166777202E-114	Endocrine
FP101882.1	0,343	0,029	0,3311042523	2,1801045667421E-113	Endocrine
anp32a	0,62	0,226	0,3341866218	7,1638860219446E-113	Endocrine
gng7	0,421	0,076	0,3030255015	7,2393861633786E-113	Endocrine
cox8a	0,834	0,406	0,3370163139	8,1258080323404E-113	Endocrine
dnajc5aa	0,322	0,017	0,2918753623	1,1080189640774E-112	Endocrine
map1lc3b	0,543	0,165	0,3145823559	1,3925214642247E-112	Endocrine
atp6v1h	0,382	0,054	0,3061094901	2,7168803230794E-112	Endocrine
sigmar1	0,423	0,085	0,2906146155	6,2904594655327E-112	Endocrine
elof1	0,569	0,188	0,3098808906	7,5171677602397E-112	Endocrine
rnf144ab	0,335	0,026	0,3298097863	9,11670003686508E-112	Endocrine
tmed3	0,331	0,025	0,3110490937	2,35922489261078E-111	Endocrine
trappc1	0,515	0,15	0,331872277	2,54792980493728E-111	Endocrine
grn1	0,36	0,045	0,4412459434	9,20030215083507E-111	Endocrine
sub1b	0,527	0,161	0,3188142745	6,3402464445829E-110	Endocrine
fkbp1aa	0,835	0,437	0,3543264476	1,5635598032797E-109	Endocrine
acbd7	0,333	0,025	0,2813642666	1,7357955865747E-108	Endocrine

lin7a	0,297	0,01	0,3092679433	3,2700201334948E-106	Endocrine
med30	0,423	0,1	0,3079392892	1,0458109109573E-105	Endocrine
tefa	0,51	0,146	0,299498699	6,773908275373E-105	Endocrine
gstk1	0,597	0,219	0,314412568	4,0906931542101E-104	Endocrine
hmgb1a	0,844	0,463	0,3679671832	6,4671599683088E-104	Endocrine
fam32a	0,742	0,364	0,3083449144	4,2131458503987E-103	Endocrine
necap1	0,389	0,067	0,2653494332	4,314887570123E-103	Endocrine
cplx2	0,297	0,013	0,3033951269	6,0222433501695E-103	Endocrine
ppme1	0,414	0,091	0,2913303607	2,1984259613515E-102	Endocrine
oaz2b	0,289	0,011	0,2929263746	6,965222490789E-102	Endocrine
atp5mc1	0,861	0,531	0,2951245315	9,9202583667051E-102	Endocrine
sox4a.1	0,4	0,074	0,3189257832	3,66719366691111E-101	Endocrine
stk35	0,318	0,028	0,3160079552	5,8432205805545E-101	Endocrine
tcp11l2	0,322	0,032	0,2931748336	5,134249756774E-100	Endocrine
krtcap2	0,696	0,352	0,3018978333	1,01387122756925E-99	Endocrine
gb:bc139872	0,336	0,044	0,2819464322	9,68442341105481E-99	Endocrine
atp5mea	0,468	0,15	0,3046717919	1,75665426403051E-98	Endocrine
si:dkey-188i13.11	0,532	0,199	0,2892754799	4,97864807197451E-98	Endocrine
si:dkey-153k10.9	0,294	0,018	0,2855618582	7,5709192477802E-98	Endocrine
cnpy1	0,584	0,26	0,3332654465	2,7481865094996E-97	Endocrine
otulina.1	0,328	0,042	0,3136271031	3,35276985187093E-96	Endocrine
nrsn1	0,28	0,013	0,2565666151	4,93411500697093E-96	Endocrine
nsfa	0,284	0,014	0,2625339394	5,06860326475108E-96	Endocrine
si:ch211-196l7.4	0,843	0,515	0,3164756575	7,76097248068376E-96	Endocrine
paip2b	0,752	0,365	0,2907062667	1,61176482118143E-95	Endocrine
znrf1	0,295	0,023	0,2598138854	1,86046491382178E-95	Endocrine
selenom	0,356	0,063	0,2637685902	5,7489739405296E-95	Endocrine
calm2a	0,361	0,064	0,2763913662	1,00057988583283E-94	Endocrine
smim7	0,372	0,074	0,2500837867	1,27724254385755E-93	Endocrine
srpr	0,564	0,226	0,2794307071	1,51045265907397E-93	Endocrine
slc16a9b	0,279	0,015	0,2993027753	1,78685082011386E-93	Endocrine
glrx	0,656	0,286	0,2701630732	3,22720324947445E-93	Endocrine
si:ch211-255i3.4	0,285	0,018	0,2645420927	7,79858748351101E-93	Endocrine
bhlhe40	0,518	0,172	0,2559985936	2,32813944614507E-92	Endocrine
phyhiplb	0,267	0,01	0,2765284081	7,85458092358207E-92	Endocrine
dpp7	0,418	0,105	0,2660700089	4,87291939436171E-91	Endocrine
anapc16	0,468	0,149	0,2556719215	2,19366801676396E-90	Endocrine
rab1ab	0,561	0,219	0,2651962283	2,40281879261527E-90	Endocrine
crebzf	0,62	0,251	0,2669325481	2,81836528254606E-90	Endocrine
cwc25	0,424	0,117	2,8225107339	2,6080068957506E-89	Endocrine
si:ch73-1a9.3	0,796	0,385	0,2858612674	1,76298218717498E-88	Endocrine
lamtor2	0,404	0,101	0,2520296408	3,74426379567975E-88	Endocrine
large2	0,3	0,034	0,2599469452	1,70185241120116E-87	Endocrine
hepacam2	0,277	0,023	0,3352311917	5,17952788952738E-86	Endocrine
clta	0,466	0,149	0,2556883978	2,30461972201974E-85	Endocrine
mllt11	0,251	0,011	0,3074738102	1,52903836084635E-84	Endocrine
sdhc	0,595	0,262	0,2570365418	8,6288136944611E-84	Endocrine
ptp4a3	0,334	0,072	0,2567504842	9,94278435893196E-83	Endocrine
fam133b	0,579	0,275	0,2733882022	4,93746874321707E-82	Endocrine
dnajc1	0,518	0,224	0,2585978607	9,02157149886234E-81	Endocrine
dedd1	0,453	0,149	0,2610175025	4,40157665515387E-79	Endocrine
rab11ba	0,442	0,153	0,2565779738	2,14893001725024E-76	Endocrine
spcs1	0,731	0,458	0,280912815	1,05917476636482E-74	Endocrine

gpx3	0,385	0,129	0,2610578438	5,03758937806654E-72	Endocrine
arrdc3a	0,302	0,061	0,2607545962	7,6519286569879E-71	Endocrine
pabpc1b	0,509	0,233	0,2559588206	6,76592296196229E-70	Endocrine
fosab	0,99	0,895	0,2889837902	2,44359815084663E-67	Endocrine
slc16a12b	0,252	0,037	0,2871807795	5,78949947606009E-67	Endocrine
calm2b	0,608	0,29	0,2691606303	1,81868606951738E-65	Endocrine
TPM1 (1 of many)	0,253	0,068	0,2683971932	1,87243483685644E-47	Endocrine
gadd45ga	0,339	0,143	0,4316751623	2,13791637163742E-44	Endocrine
nupr1	0,916	0,318	1,7840820229	1,1561174885165E-252	Endocrine
ponzr1	0,807	0,333	0,4115996631	1,7784351625359E-101	Endocrine
npm1a	0,807	0,381	0,4924225077	2,396403758034E-93	Endocrine
eef2b	0,951	0,77	0,4118810963	2,03704052936308E-67	Endocrine
rpl12	0,998	0,994	0,3266345536	2,12354215527814E-66	Endocrine
aldob	0,906	0,606	0,3705595969	2,84059484408985E-51	Endocrine
prdx4	0,911	0,683	0,3019215565	1,07830923168051E-49	Endocrine
pklr	0,725	0,378	0,2687534689	2,10705230638528E-46	Endocrine
anxa4	0,663	0,405	0,272882826	1,2841210511134E-14	Endocrine
tmsb1	0,998	0,792	0,9667243519	9.99988867182683e-321	Endocrine
pdx1	0,893	0,305	0,3398122405	3,6579523197094E-237	Endocrine
krt18	0,666	0,283	0,5948089191	1,1922587007788E-162	Endocrine
hsp90aa1.2	0,999	0,992	0,5061586714	2,6209608401995E-158	Endocrine
rpl13a	1	0,991	0,2731490625	1,1996239841683E-134	Endocrine
rpl5a	0,991	0,934	0,2951388364	4,1794641891832E-114	Endocrine
hspe1	0,96	0,804	0,3816453738	1,1337805444727E-113	Endocrine
rpl14	0,993	0,944	0,2691638148	5,4872322417004E-109	Endocrine
zgc:92066	0,901	0,745	0,2871577941	2,69482621817269E-64	Endocrine
rps27.1	0,999	0,997	0,6231980487	1,8508343723031E-250	Endocrine
b2m	0,991	0,964	1,0431018501	1,6812744946152E-238	Endocrine
slc25a5	1	0,993	0,6467302774	6,3317660537517E-237	Endocrine
mt-co1	0,999	0,99	0,7120631666	8,1101386031921E-236	Endocrine
naca	0,999	0,985	0,648370628	6,016368727571E-210	Endocrine
rpl29	0,981	0,967	0,6891681232	7,85681105121174E-200	Endocrine
mt-nd4	0,968	0,913	0,7773973519	2,1516106594218E-192	Endocrine
mt-nd1	0,937	0,781	0,8325265511	2,853502685913E-178	Endocrine
zgc:171772	1	0,996	0,5602750344	1,2597184284147E-171	Endocrine
rps21	0,993	0,978	0,6473282156	8,6019986307014E-170	Endocrine
mt-co2	1	0,999	0,4537950602	7,7386680264347E-133	Endocrine
mt-nd2	0,929	0,81	0,6140364695	6,247351118981E-113	Endocrine
romo1	0,92	0,818	0,5929272782	4,8848075334622E-104	Endocrine
selenow1	0,878	0,804	0,4831612516	5,27395821403322E-69	Endocrine
actb2	0,966	0,919	0,4403768298	3,45435165789356E-65	Endocrine
ccni	0,81	0,707	0,434060116	1,47143881061555E-52	Endocrine
mt-nd3	0,822	0,759	0,3995027305	6,68520502712465E-39	Endocrine
ran	0,882	0,836	0,2971869636	7,43927971007101E-28	Endocrine
cirbpa	0,834	0,835	0,2751191858	4,69836323268849E-15	Endocrine
h3f3b.1	0,621	0,528	0,2889228428	6,28723906848707E-15	Endocrine
zgc:162730	0,832	0,471	0,8755178587	9.99988867182683e-321	Endocrine
cebpb	0,675	0,293	0,8367961185	1,6509970509145E-213	Endocrine
btg2	0,936	0,795	0,4980649651	1,2483105258007E-183	Endocrine
nfkbiab	0,826	0,569	0,5038267119	1,4375131146645E-154	Endocrine
cox8b	0,537	0,226	0,3625294417	8,3623932812256E-123	Endocrine
mdh1aa	0,809	0,704	0,4209651599	3,1253236989972E-117	Endocrine
glula	0,535	0,241	0,3771465379	3,8759991984636E-117	Endocrine

egr1	0,54	0,303	0,2854283676	1,58939008078539E-69	Endocrine
lgals2b	0,993	0,058	4,7354754106	9.99988867182683e-321	Ductal
zgc:193726	0,75	0,032	4,2141242589	9.99988867182683e-321	Ductal
phlda2	0,895	0,095	3,4806889093	9.99988867182683e-321	Ductal
si:dkey-96g2.1	0,963	0,012	3,1664138031	9.99988867182683e-321	Ductal
cnn2	0,902	0,032	2,8916973593	9.99988867182683e-321	Ductal
si:ch211-153b23.5	0,749	0,06	2,8830519446	9.99988867182683e-321	Ductal
cfl1l	0,682	0,136	2,6857546924	9.99988867182683e-321	Ductal
BX855618.1	0,531	0,014	2,6634639513	9.99988867182683e-321	Ductal
cldn15a	0,897	0,083	2,5784907482	9.99988867182683e-321	Ductal
cldnc	0,944	0,108	2,4290335256	9.99988867182683e-321	Ductal
cd9b	0,889	0,01	2,4280055355	9.99988867182683e-321	Ductal
tagln2	0,826	0,11	2,3766663984	9.99988867182683e-321	Ductal
si:ch73-335l21.4	0,801	0,029	2,3728117761	9.99988867182683e-321	Ductal
hbegfa	0,841	0,115	2,343787887	9.99988867182683e-321	Ductal
f3b	0,76	0,182	2,2504374158	9.99988867182683e-321	Ductal
zgc:158343	0,689	0,041	2,2406693142	9.99988867182683e-321	Ductal
elf3	0,842	0,045	2,228300206	9.99988867182683e-321	Ductal
s100a10b	0,911	0,259	2,2221787335	9.99988867182683e-321	Ductal
her9	0,674	0,065	2,1731175367	9.99988867182683e-321	Ductal
krt94	0,555	0,037	2,1645131086	9.99988867182683e-321	Ductal
ccl19a.1	0,749	0,024	2,162425685	9.99988867182683e-321	Ductal
cldn7b	0,782	0,037	2,1491790766	9.99988867182683e-321	Ductal
spint2	0,912	0,191	2,0748153091	9.99988867182683e-321	Ductal
gpx1b	0,767	0,032	2,0227309424	9.99988867182683e-321	Ductal
jdp2b	0,759	0,096	1,9969955534	9.99988867182683e-321	Ductal
fosl1a	0,821	0,266	1,9898756564	9.99988867182683e-321	Ductal
dusp5	0,738	0,101	1,9677255871	9.99988867182683e-321	Ductal
her6	0,651	0,096	1,9617304634	9.99988867182683e-321	Ductal
pfn1	0,867	0,204	1,9382429125	9.99988867182683e-321	Ductal
hmox1a	0,687	0,011	1,9085683513	9.99988867182683e-321	Ductal
ctgfa	0,629	0,007	1,9048457673	9.99988867182683e-321	Ductal
ftr82	0,803	0,012	1,8909585004	9.99988867182683e-321	Ductal
cdh1	0,817	0,055	1,8629659131	9.99988867182683e-321	Ductal
krt4	0,554	0,017	1,8312289849	9.99988867182683e-321	Ductal
zgc:64022	0,535	0,014	1,7355757031	9.99988867182683e-321	Ductal
fam213b	0,754	0,093	1,7097460621	9.99988867182683e-321	Ductal
hnrnpa0l.1	0,933	0,728	1,6963998256	9.99988867182683e-321	Ductal
rpz5	0,741	0,195	1,6884556923	9.99988867182683e-321	Ductal
efna1b	0,633	0,012	1,682879957	9.99988867182683e-321	Ductal
myl9b	0,783	0,136	1,6814837539	9.99988867182683e-321	Ductal
f11r.1	0,87	0,435	1,6783324327	9.99988867182683e-321	Ductal
sdc4	0,883	0,543	1,6767276117	9.99988867182683e-321	Ductal
btg1	0,984	0,901	1,6137830781	9.99988867182683e-321	Ductal
cyr61	0,648	0,071	1,5982111124	9.99988867182683e-321	Ductal
pdlim1	0,771	0,058	1,5944736504	9.99988867182683e-321	Ductal
mych	0,679	0,099	1,5788776086	9.99988867182683e-321	Ductal
krt91	0,515	0,013	1,5644332792	9.99988867182683e-321	Ductal
rassf7b	0,755	0,055	1,557165847	9.99988867182683e-321	Ductal
klf6a	0,691	0,077	1,5439882206	9.99988867182683e-321	Ductal
serpinb1	0,797	0,259	1,5370950831	9.99988867182683e-321	Ductal
tpm3	0,774	0,181	1,5295037621	9.99988867182683e-321	Ductal
cd74a	0,506	0,033	1,5198118171	9.99988867182683e-321	Ductal

lxn	0,682	0,107	1,5095604005	9.99988867182683e-321	Ductal
arl5c	0,675	0,098	1,499078573	9.99988867182683e-321	Ductal
id2a	0,67	0,06	1,4943265354	9.99988867182683e-321	Ductal
zgc:198329	0,715	0,013	1,4939284651	9.99988867182683e-321	Ductal
selenop	0,723	0,165	1,4918272024	9.99988867182683e-321	Ductal
si:ch73-86n18.1	0,61	0,02	1,4896901333	9.99988867182683e-321	Ductal
ccl19b	0,587	0,005	1,4692936929	9.99988867182683e-321	Ductal
adh8b	0,722	0,127	1,4055643999	9.99988867182683e-321	Ductal
sgk1	0,654	0,059	1,3864609044	9.99988867182683e-321	Ductal
s100v2	0,682	0,016	1,3800799471	9.99988867182683e-321	Ductal
sb:cb1058	0,673	0,038	1,3634481629	9.99988867182683e-321	Ductal
cx30.3	0,595	0,005	1,2658810833	9.99988867182683e-321	Ductal
si:dkey-16l2.20	0,542	0,004	1,2260542925	9.99988867182683e-321	Ductal
cdh17	0,565	0,004	1,1298734399	9.99988867182683e-321	Ductal
zgc:92313	0,643	0,123	1,1176450549	9.99988867182683e-321	Ductal
trib3	0,593	0,095	1,1055040654	9.99988867182683e-321	Ductal
spint1a	0,607	0,051	1,0696087802	9.99988867182683e-321	Ductal
tspan35	0,52	0,023	1,0675943612	9.99988867182683e-321	Ductal
cbln17	0,529	0,05	1,0533991827	9.99988867182683e-321	Ductal
dhrs9	0,519	0,025	0,9832081984	9.99988867182683e-321	Ductal
perp	0,529	0,05	0,9636086566	9.99988867182683e-321	Ductal
zfp36l2	0,523	0,036	0,9570733954	9.99988867182683e-321	Ductal
lrrfip1a	0,547	0,096	0,9413287555	9,380158406709E-298	Ductal
pdzk1ip1	0,686	0,201	1,1730614663	1,637385677853E-294	Ductal
rhoab	0,651	0,174	1,3124018648	3,5379955806393E-293	Ductal
ecm1b	0,549	0,1	1,0748198863	1,4049974311381E-289	Ductal
prdx2	0,84	0,503	1,4831330549	8,1092596643588E-287	Ductal
si:dkey-180p18.9	0,509	0,085	0,8282826496	2,1170854479804E-280	Ductal
chac1	0,865	0,613	1,6346788648	5,9410172165703E-276	Ductal
tubb4b	0,894	0,713	1,243469769	4,2877190466417E-274	Ductal
gadd45bb	0,666	0,199	1,6315807611	4,4446972803214E-270	Ductal
eno3	0,715	0,272	1,2669325864	9,045599711595E-263	Ductal
eif1b	0,949	0,913	1,0128817639	6,8823557175272E-261	Ductal
tuba8l	0,697	0,26	1,2285265442	8,3983090157847E-256	Ductal
tpm4a	0,557	0,123	1,0925390855	2,2971957458881E-255	Ductal
zgc:109934	0,624	0,179	1,1079302887	3,2051504228094E-254	Ductal
socs3a	0,578	0,136	1,0083814648	4,7699790682418E-254	Ductal
maff	0,661	0,221	1,2352787771	4,1009068090584E-251	Ductal
jupa	0,518	0,104	0,8964654705	2,6407663531447E-250	Ductal
serinc2	0,585	0,148	0,9208009589	1,7909900864657E-249	Ductal
oclna	0,582	0,158	1,052390964	6,5287144311961E-235	Ductal
tob1b	0,854	0,713	1,2849434939	9,5243804108216E-230	Ductal
zfand5a	0,61	0,196	1,0599139002	5,0189446044514E-222	Ductal
cotl1	0,653	0,249	1,3306276575	3,8762195649406E-221	Ductal
mycb	0,623	0,211	1,1515498291	5,9144672288371E-212	Ductal
foxp4	0,682	0,293	1,140100816	2,6970654395154E-209	Ductal
ppp1r10	0,673	0,29	1,1812507539	1,5097546927301E-204	Ductal
glulb	0,606	0,208	1,0738370521	3,5365712919379E-204	Ductal
prelid3b	0,815	0,621	1,1954970066	1,1869114234491E-200	Ductal
dap1b	0,674	0,309	1,1213379331	7,1387799318622E-197	Ductal
alas1	0,51	0,14	0,9807181068	8,3736033396274E-195	Ductal
cox6a2	0,695	0,333	1,0890139052	1,967347322222E-191	Ductal
zgc:153867	0,818	0,68	1,0080534869	4,0429312699564E-191	Ductal

gstr	0,705	0,369	1,0465478009	8,8054907307037E-187	Ductal
arg2	0,586	0,215	1,43459626	4,6569402001972E-185	Ductal
fosl2	0,711	0,377	1,1135518465	1,1038369441432E-183	Ductal
pfn2	0,79	0,582	0,9826442191	1,4563767132817E-182	Ductal
eif4ebp1	0,642	0,291	1,2838385771	1,6136604144782E-182	Ductal
brdt	0,544	0,172	0,8970994806	6,1707475996236E-182	Ductal
tspo	0,514	0,159	0,8534937787	6,2441006297341E-174	Ductal
brd2a	0,665	0,331	1,043449735	1,2322804036149E-172	Ductal
slc25a25a	0,582	0,23	1,0146334748	9,092000843631E-171	Ductal
atf4b	0,816	0,703	1,0099841937	2,1779982452018E-167	Ductal
cycsb	0,781	0,609	1,1293247095	6,9326946708985E-167	Ductal
slc25a55a	0,603	0,247	0,8982725226	8,9828252738925E-166	Ductal
nocta	0,662	0,353	1,0806893797	3,7890530402339E-155	Ductal
arpc2	0,656	0,383	0,870553623	2,8633023089315E-139	Ductal
cxcl20	0,552	0,233	0,9235134583	2,9239171081939E-136	Ductal
tmem176l.1	0,552	0,244	0,78595103	2,1599029321449E-131	Ductal
basp1	0,503	0,183	0,7874294043	8,9796796588714E-131	Ductal
ccng1	0,777	0,693	0,9763280539	2,5155947570895E-124	Ductal
itm2bb	0,531	0,228	0,7661944572	1,4209259752065E-123	Ductal
aldh9a1a.1	0,558	0,269	0,776183044	5,2679158073466E-123	Ductal
si:ch211-212k18.7	0,552	0,247	0,7373870038	1,5433435035443E-122	Ductal
elovl1b	0,694	0,482	0,7922756483	1,463546891886E-117	Ductal
ddit3	0,575	0,283	0,8862543198	7,6338873793322E-115	Ductal
ncl	0,655	0,418	0,8133575213	4,4373089889263E-114	Ductal
zgc:92027	0,504	0,224	0,6552795558	2,816148160499E-102	Ductal
ets2	0,551	0,291	0,818200236	7,9882009020166E-102	Ductal
uqcrb	0,666	0,516	0,6944564221	2,38026189151146E-90	Ductal
hnrnpa1b	0,599	0,385	0,6803705747	2,38129692215463E-83	Ductal
ip6k2b	0,667	0,533	0,8403093264	9,26594653259439E-83	Ductal
ak2	0,615	0,44	0,6978902403	1,07529159985629E-82	Ductal
atp5f1b	0,873	0,863	0,4962685557	7,2970273346929E-82	Ductal
pdia3	0,789	0,797	0,6133503977	6,45802012898519E-81	Ductal
arpc1a	0,538	0,309	0,6537319391	5,52845607076961E-80	Ductal
si:dkey-177p2.6	0,722	0,679	0,7904609467	6,63392488065705E-78	Ductal
ranbp1	0,597	0,419	0,6741677728	6,15228648039056E-75	Ductal
rca2.1	0,545	0,349	0,6912319459	1,18886381543151E-70	Ductal
ahsa1b	0,609	0,442	0,768989361	5,28549842472573E-69	Ductal
erh	0,792	0,79	0,5596520398	5,97394965366622E-69	Ductal
atp5if1a	0,598	0,458	0,6848116	2,39273296875626E-66	Ductal
arpc3	0,513	0,316	0,6418136455	6,27274414682996E-66	Ductal
psap	0,535	0,351	0,6160778581	3,32484072649068E-64	Ductal
bsg	0,738	0,699	0,5156538649	2,08635460669714E-62	Ductal
stip1	0,654	0,547	0,7730506273	4,40449399224995E-59	Ductal
zfand5b	0,586	0,457	0,7209892993	6,57266533697229E-56	Ductal
aldoaa	0,57	0,429	0,6049088831	3,52244585424201E-54	Ductal
hspbp1	0,746	0,675	0,7727691638	7,93516072550358E-53	Ductal
tsc22d3	0,569	0,425	0,6514068441	8,71128017814074E-52	Ductal
CCT2	0,546	0,41	0,5761055424	3,83418136808866E-50	Ductal
prdx6	0,541	0,405	0,5501181326	1,09684199070943E-48	Ductal
si:dkey-7j14.6	0,603	0,51	0,541308195	4,94954075723444E-47	Ductal
tcp1	0,555	0,434	0,5754282586	1,62461092984967E-46	Ductal
tomm5	0,574	0,461	0,5342121894	1,98983421308803E-45	Ductal
seta	0,565	0,449	0,539827659	3,71556945804348E-44	Ductal

eif3k	0,565	0,458	0,5694568767	3,89596922381467E-43	Ductal
ndfip1	0,513	0,371	0,4940551571	1,5141515634233E-42	Ductal
psma1	0,529	0,398	0,5180329821	1,57381791857066E-42	Ductal
bag3	0,649	0,544	0,607121634	5,29410975951538E-38	Ductal
hnrnpabb	0,581	0,495	0,5374098869	6,09705317205249E-38	Ductal
cct3	0,575	0,498	0,4810451923	9,83080842086823E-37	Ductal
tuba8l4	0,614	0,554	0,5424480457	2,10763522635417E-36	Ductal
cct4	0,569	0,515	0,5334310122	1,007962804729E-35	Ductal
cnbpb	0,691	0,682	0,439481148	1,21999262879867E-35	Ductal
serf2	0,737	0,777	0,4121928176	1,0323266041155E-34	Ductal
atp5fa1	0,645	0,623	0,4781972776	4,32287754521286E-33	Ductal
rbm4.3	0,584	0,529	0,5228109799	9,96602237544173E-32	Ductal
atp5mf	0,616	0,619	0,5086287828	1,22570965471005E-29	Ductal
setb	0,523	0,446	0,4882508051	9,26238704679112E-28	Ductal
uqcrh	0,64	0,677	0,4135271915	8,34986517852917E-25	Ductal
cct5	0,567	0,555	0,4217825814	9,11771160887317E-19	Ductal
pgrmc1	0,53	0,513	0,3980081151	2,56441781640847E-14	Ductal
ldhba	0,568	0,582	0,4097358901	3,95981937636002E-14	Ductal
ywhaba	0,521	0,502	0,380066492	2,4405055093968E-12	Ductal
hmgn7	0,653	0,705	0,2801099746	8,13595437104151E-09	Ductal
ywhabb	0,522	0,545	0,3136010184	0,000004005	Ductal
ywhae1	0,559	0,605	0,2846180127	0,0007994639	Ductal
spink4	0,987	0,291	4,3052089436	9.99988867182683e-321	Acinar
ctrl	0,998	0,676	4,203908781	9.99988867182683e-321	Acinar
si:ch211-240l19.8	0,998	0,837	3,8546715477	9.99988867182683e-321	Acinar
cpa5	0,998	0,587	3,846381946	9.99988867182683e-321	Acinar
CELA1 (1 of many).1	0,995	0,372	3,8233267606	9.99988867182683e-321	Acinar
cel.1	0,991	0,594	3,748973483	9.99988867182683e-321	Acinar
amy2a	0,995	0,364	3,7204132208	9.99988867182683e-321	Acinar
zgc:92590	0,998	0,651	3,65028612	9.99988867182683e-321	Acinar
cel.2	0,995	0,407	3,624424639	9.99988867182683e-321	Acinar
si:ch211-240l19.6	0,993	0,603	3,5877116138	9.99988867182683e-321	Acinar
si:ch211-240l19.7	0,984	0,277	3,50764717	9.99988867182683e-321	Acinar
cpa1	0,991	0,244	3,4542659925	9.99988867182683e-321	Acinar
si:ch211-240l19.5	0,996	0,54	3,43130585	9.99988867182683e-321	Acinar
zgc:112368	0,984	0,213	3,151342391	9.99988867182683e-321	Acinar
si:dkey-14d8.6	0,991	0,382	3,1473839222	9.99988867182683e-321	Acinar
sycn.2	0,991	0,394	3,1146479782	9.99988867182683e-321	Acinar
c6ast3	0,993	0,399	3,0482208009	9.99988867182683e-321	Acinar
zgc:136461	0,954	0,221	2,9919195279	9.99988867182683e-321	Acinar
si:dkey-14d8.7	0,982	0,223	2,9330275856	9.99988867182683e-321	Acinar
pla2g1b	0,976	0,173	2,8579912602	9.99988867182683e-321	Acinar
sycn.1	0,984	0,446	2,8463031313	9.99988867182683e-321	Acinar
pglyrp6	0,934	0,065	2,8393861581	9.99988867182683e-321	Acinar
zgc:92137	0,952	0,052	2,5335246051	9.99988867182683e-321	Acinar
CELA1 (1 of many).4	0,938	0,08	2,4814513293	9.99988867182683e-321	Acinar
pdia2	0,971	0,061	2,3970504739	9.99988867182683e-321	Acinar
si:ch211-195b11.3	0,954	0,125	2,2972844007	9.99988867182683e-321	Acinar
CT027638.1	0,969	0,35	2,2531787278	9.99988867182683e-321	Acinar
CELA1 (1 of many)	0,941	0,11	2,2073210873	9.99988867182683e-321	Acinar
endou	0,941	0,082	2,0424565282	9.99988867182683e-321	Acinar
cpa2	0,936	0,057	2,0390823598	9.99988867182683e-321	Acinar
c6ast4	0,931	0,125	2,0266574803	9.99988867182683e-321	Acinar

cpa4	0,932	0,053	2,0145444248	9.99988867182683e-321	Acinar
gamt	0,952	0,22	1,8069031817	9.99988867182683e-321	Acinar
CR753876.1	0,846	0,017	1,6779671966	9.99988867182683e-321	Acinar
CELA1 (1 of many).3	0,867	0,138	1,6256795722	9.99988867182683e-321	Acinar
dnase1	0,879	0,024	1,5986074131	9.99988867182683e-321	Acinar
bhmt	0,91	0,087	1,4992861571	9.99988867182683e-321	Acinar
aqp12	0,892	0,107	1,453926951	9.99988867182683e-321	Acinar
ahcy	0,907	0,21	1,4027495294	9.99988867182683e-321	Acinar
AL845362.1	0,636	0,011	0,7797172858	9.99988867182683e-321	Acinar
pck1	0,523	0,031	0,6224195392	9.99988867182683e-321	Acinar
cox6b1	0,589	0,057	0,5851435761	9.99988867182683e-321	Acinar
pabpc4	0,506	0,004	0,5788997295	9.99988867182683e-321	Acinar
mat1a	0,565	0,016	0,5649800221	9.99988867182683e-321	Acinar
gapdh	0,75	0,117	0,6660830228	1,8554010176391E-272	Acinar
cox5b2	0,548	0,069	0,4745801007	2,9611195595337E-253	Acinar
wu:fj16a03	0,583	0,091	0,8602536463	8,8093316733437E-226	Acinar
hdlbpa	0,799	0,257	0,8607007973	3,2579298841348E-208	Acinar
apoa1b	0,589	0,115	0,8237720966	1,1794381654546E-184	Acinar
tfa	0,548	0,096	0,7163309348	3,1950551837912E-184	Acinar
fabp10a	0,627	0,162	1,2411683247	6,567963524787E-166	Acinar
apoc1	0,598	0,144	0,9622399008	2,0913975862847E-163	Acinar
fkbp11	0,695	0,229	0,6896182262	2,2265318156437E-160	Acinar
SERP1	0,951	0,892	0,6329263308	1,4331460631943E-145	Acinar
apoa2	0,638	0,253	1,1557426059	2,94286064588827E-111	Acinar
snd1	0,625	0,227	0,4912524407	4,5174336916398E-100	Acinar
sec61a1	0,693	0,33	0,4641493134	8,91855878917427E-81	Acinar
dap	0,742	0,462	0,5320164537	2,78745805016103E-74	Acinar
idh2	0,691	0,353	0,5499041597	1,50125972046268E-68	Acinar
eef1da	0,837	0,591	0,3008538073	5,57633450961463E-51	Acinar
ndufa6	0,651	0,355	0,3350414418	1,44448331502174E-43	Acinar

Supplemental table 3

Set of significative marker genes (pvalue adjusted <= 0.1) identified for all cell types based on the integration of human scRNA-seq datasets.

Markers SC Human

Gene	Pct.Cluster	Pct.Others	Avg_LogFC	Pval_Adjusted	Cluster
CELA2A	0,963	0,118	3,69159456	9.99988867182683e-321	Acinar
CELA2B	0,78	0,025	1,8007152793	9.99988867182683e-321	Acinar
CELA3A	0,976	0,287	4,4743024454	9.99988867182683e-321	Acinar
CELA3B	0,933	0,155	4,0764600663	9.99988867182683e-321	Acinar
CLPS	0,92	0,211	4,1457941019	9.99988867182683e-321	Acinar
CPA1	0,987	0,251	4,0452616515	9.99988867182683e-321	Acinar
CPA2	0,988	0,153	3,2803373415	9.99988867182683e-321	Acinar
CTRB1	0,995	0,284	4,2425369446	9.99988867182683e-321	Acinar
CTRB2	0,999	0,366	4,0929445974	9.99988867182683e-321	Acinar
CTRC	0,983	0,125	3,2211106212	9.99988867182683e-321	Acinar
CTRL	0,809	0,03	1,3833288441	9.99988867182683e-321	Acinar
GP2	0,863	0,113	2,3914784187	9.99988867182683e-321	Acinar
KLK1	0,951	0,046	2,1774499439	9.99988867182683e-321	Acinar
MGST1	0,923	0,157	1,2610771754	9.99988867182683e-321	Acinar
PLA2G1B	0,989	0,21	3,8668824169	9.99988867182683e-321	Acinar
PNLIP	0,961	0,17	3,4338669303	9.99988867182683e-321	Acinar
PNLIPRP1	0,946	0,1	2,7976533689	9.99988867182683e-321	Acinar
PNLIPRP2	0,919	0,103	2,4152708968	9.99988867182683e-321	Acinar
PRSS1	0,999	0,331	4,2286426217	9.99988867182683e-321	Acinar
PRSS3	0,972	0,292	2,6489707268	9.99988867182683e-321	Acinar
REG1A	0,999	0,654	3,9063561889	9.99988867182683e-321	Acinar
REG1B	0,974	0,335	3,8164305772	9.99988867182683e-321	Acinar
RNASE1	0,909	0,133	1,8558209749	9.99988867182683e-321	Acinar
SYCN	0,769	0,041	2,9593649178	9.99988867182683e-321	Acinar
UGDH	0,965	0,475	1,4144538002	9.99988867182683e-321	Acinar
EEF1A1	0,999	0,997	0,5969796523	1,23255631968705E-247	Acinar
REG1P	0,634	0,01	1,2245771466	9,83349835711324E-243	Acinar
PLTP	0,683	0,049	0,9258613332	4,17621661963916E-241	Acinar
REG3G	0,701	0,08	2,3455562064	4,24287055644744E-240	Acinar
CUZD1	0,637	0,024	2,3378752451	7,67540939214768E-236	Acinar
ANPEP	0,877	0,281	1,2261890784	1,40571351787741E-225	Acinar
PRSS2	0,758	0,247	4,331155769	1,51620343471734E-223	Acinar
ALDOB	0,74	0,143	1,6654328824	1,99710858340036E-223	Acinar
PDIA2	0,575	0,021	0,7761141138	5,15551277743122E-204	Acinar
BCAT1	0,617	0,056	1,1718878504	1,01246434866152E-201	Acinar
AMY2A	0,569	0,023	1,5023079238	1,66682595152041E-201	Acinar
GSTA2	0,657	0,121	1,5887505357	2,17952958146009E-190	Acinar
RPS2	0,997	0,989	0,6346684294	1,78726913410873E-185	Acinar
CPB1	0,985	0,387	2,968904131	5,30850443709444E-173	Acinar
FAM129A	0,55	0,061	0,6023884731	1,78429564866161E-158	Acinar
RPS5	0,999	0,971	0,6032552259	3,51058965182765E-150	Acinar
GAS5	0,954	0,686	0,756148844	1,54895162130941E-149	Acinar
CXCL17	0,548	0,081	0,8128846979	2,65230694076186E-145	Acinar
RARRES2	0,552	0,091	0,8854110377	7,80410520009702E-144	Acinar
RPL10	0,999	0,979	0,8303577904	3,08576003948484E-140	Acinar
ALB	0,537	0,109	1,9153073951	2,72487074955642E-135	Acinar
EEF1B2	0,995	0,913	0,6755287445	4,81171840307148E-134	Acinar
RPL12	0,995	0,896	0,6554967974	2,78206944031423E-131	Acinar
RPS6	1	0,992	0,3398338038	4,53987068117679E-129	Acinar
IL32	0,988	0,828	0,8907888012	3,3539413000264E-123	Acinar
RPL13	1	0,996	0,3723836111	1,88556867830189E-120	Acinar
FTH1	1	0,994	0,5519495588	2,25542983954903E-119	Acinar

SERPINB1	0,946	0,697	0,700875189	3,13277972767638E-114	Acinar
RPL35A	1	0,976	0,406020308	3,2021441766415E-111	Acinar
CEL	0,721	0,063	2,0209896629	3,38668192073684E-104	Acinar
GSTA1	0,618	0,244	1,30480373	4,10035191279992E-101	Acinar
RPLP0	1	0,983	0,5345026808	9,68763851243476E-100	Acinar
TRIB2	0,594	0,194	0,4983695287	1,98080237335674E-99	Acinar
F3	0,656	0,231	0,5517757875	7,30586683214951E-97	Acinar
SERPINA3	0,993	0,811	0,7879671359	1,02163253057658E-95	Acinar
RPL32	1	0,982	0,3650028394	4,7355973948119E-95	Acinar
RPL14	0,997	0,969	0,430689301	1,65615714429525E-94	Acinar
GNB2L1	1	0,967	0,4142959841	1,73207888553804E-94	Acinar
MT1G	0,562	0,192	1,6522486258	1,12827564007061E-93	Acinar
RPS4Y1	0,832	0,528	0,441351847	1,00458462962749E-87	Acinar
PABPC4	0,858	0,554	0,4898276532	1,9857709346736E-84	Acinar
RPS12	1	0,98	0,4332919701	4,61000684393515E-83	Acinar
GLTSCR2	0,936	0,64	0,3787967487	2,21136315607796E-82	Acinar
REG3A	0,565	0,274	3,154802064	6,51265986286449E-79	Acinar
IMPDH2	0,884	0,574	0,4736499028	8,5926959746403E-79	Acinar
RPL36A	0,924	0,884	0,4904836719	2,45137518625211E-74	Acinar
RPL36	0,999	0,96	0,40007721	2,13657275821817E-73	Acinar
AKR1C3	0,761	0,424	0,6432164974	5,66051007880874E-73	Acinar
RPL13A	0,999	0,994	0,3595134435	6,04444356060402E-73	Acinar
PTGR1	0,7	0,395	0,7943609869	2,39212528636405E-70	Acinar
RPL18A	0,996	0,952	0,4920161234	5,08715720861832E-70	Acinar
RPS9	0,999	0,98	0,4257392223	6,3349119764003E-69	Acinar
FKBP11	0,717	0,413	0,3017440615	1,89120184037145E-68	Acinar
TKT	0,961	0,782	0,4589731328	8,87802574108057E-67	Acinar
RPS8	0,999	0,994	0,2982126204	4,28671746033305E-66	Acinar
SLC25A37	0,82	0,493	0,3814176878	6,42516806571049E-65	Acinar
SNHG5	0,93	0,731	0,5081248984	7,82319001066975E-65	Acinar
GALNT2	0,811	0,497	0,3949290386	8,61971876758814E-64	Acinar
SORD	0,612	0,291	0,3568280217	4,27920058904509E-61	Acinar
EEF1D	0,969	0,852	0,3870248816	2,1543926838258E-60	Acinar
RPS18	0,983	0,967	0,3581101666	3,71595086384794E-60	Acinar
PGM1	0,72	0,403	0,2727954375	8,48601435021901E-60	Acinar
RPL35	0,997	0,954	0,4487200758	1,24451597433405E-58	Acinar
RPL8	0,998	0,991	0,3111632901	1,41289738996855E-56	Acinar
RPL9	0,972	0,865	0,3770281524	2,43700279356701E-55	Acinar
RPL29	0,974	0,892	0,4825162195	5,09237240161323E-55	Acinar
NOB1	0,695	0,357	0,2915666398	4,24775036953302E-54	Acinar
GATM	0,813	0,526	0,5431069775	3,50110445911395E-53	Acinar
CEBPD	0,862	0,579	0,4223402001	1,29638924448768E-50	Acinar
HSPA8	0,968	0,879	0,4051364509	4,06244316879597E-50	Acinar
SRM	0,72	0,4	0,2638877596	7,36120788480655E-48	Acinar
POLR1D	0,861	0,608	0,3228834079	9,68489588868074E-48	Acinar
RPS20	0,997	0,977	0,2598253566	1,40774911271311E-47	Acinar
RPSA	0,99	0,95	0,4183427794	3,55806596149529E-47	Acinar
RPS14	0,996	0,977	0,2732596707	2,84892837351813E-45	Acinar
P4HB	0,978	0,906	0,2793313369	3,96989153831383E-42	Acinar
AGTRAP	0,702	0,433	0,4323588825	1,4736769300758E-38	Acinar
ACTG1	1	0,998	0,4859390321	2,08441350970863E-37	Acinar
SOD2	0,982	0,886	0,4204082527	1,26464885704807E-35	Acinar
HNRNPA1L2	0,793	0,536	0,2781874429	5,99168035692148E-32	Acinar

OAF	0,667	0,409	0,5179853681	5,82829788231313E-31	Acinar
C15orf48	0,61	0,415	0,480877003	1,65975914891366E-30	Acinar
CA12	0,571	0,346	0,4518125715	1,58754216043435E-29	Acinar
TC2N	0,682	0,473	0,3049931015	3,89560771205314E-29	Acinar
AKR1C2	0,516	0,303	0,6570363457	9,99065788277678E-27	Acinar
RPS26	0,904	0,835	0,2884948745	3,48973157158103E-25	Acinar
RPL13AP17	0,648	0,508	0,2709578029	5,50131092146591E-22	Acinar
GLUL	0,812	0,615	0,4032086706	1,21720539481378E-20	Acinar
CYB5A	0,792	0,67	0,272853658	8,59726925312305E-14	Acinar
VGF	0,969	0,833	0,8601649466	1,52342913952852E-215	Alpha
GLS	0,797	0,502	1,0235319312	6,49903189993781E-203	Alpha
CD99L2	0,584	0,27	0,6312123367	7,42925478984952E-194	Alpha
SPINT2	0,979	0,964	0,5067501504	1,70064637941393E-121	Alpha
SLC7A2	0,926	0,702	1,0754368618	1,57114691725894E-93	Alpha
MAFB	0,792	0,649	0,5065535368	6,27752744917625E-81	Alpha
CAMK2G	0,602	0,265	0,8732809359	3,27538518533574E-75	Alpha
PCSK2	0,965	0,92	0,8292619507	1,01928737058129E-69	Alpha
HIGD1A	0,714	0,538	0,476468529	5,90868006393044E-61	Alpha
COTL1	0,709	0,5	0,5441207854	7,27819965511867E-58	Alpha
CRYBA2	0,685	0,402	1,6007237212	2,67912717126236E-52	Alpha
ERP29	0,879	0,785	0,2854192791	8,75109305302594E-46	Alpha
GPX3	0,96	0,896	0,5203920742	1,45021877624324E-41	Alpha
FXYD5	0,622	0,322	1,0297748475	5,24313385648336E-36	Alpha
SLC7A8	0,723	0,656	0,2824751268	4,66114309458124E-26	Alpha
RGS4	0,817	0,598	0,5509317013	5,56708114302987E-24	Alpha
PDK4	0,769	0,651	0,7159825931	3,7345122924439E-14	Alpha
TTR	1	1	1,4829570483	1,49381735735136E-11	Alpha
GCG	0,998	0,769	2,4048068851	3,35686774232259E-11	Alpha
SYT5	0,514	0,384	0,3459915821	2,55059179462462E-10	Alpha
CST3	0,972	0,954	0,3367219114	2,9197475815547E-10	Alpha
ALDH1A1	0,94	0,918	0,3058801062	2,41522715846473E-06	Alpha
IRX2	0,841	0,154	1,3073102494	0,000006632	Alpha
CDKN1C	0,714	0,461	1,8745420654	2,65805894805933E-216	Beta
DLK1	0,56	0,109	2,2643672577	1,60699871633964E-180	Beta
RPL3	0,998	0,999	0,5751108556	5,83586162903751E-176	Beta
ADCYAP1	0,792	0,191	2,0893609327	7,37756694992849E-115	Beta
MEG3	0,607	0,314	0,9067306481	1,21026144487299E-106	Beta
SLC6A6	0,539	0,212	0,8606263773	2,92537220874684E-104	Beta
MAFA	0,572	0,041	1,4229506555	2,87209943975694E-76	Beta
ERO1B	0,66	0,38	1,0382310348	8,22349456444483E-74	Beta
HADH	0,883	0,724	0,5858899238	1,95912020198405E-68	Beta
ATP6V1G1	0,903	0,864	0,2961202296	2,29378925326781E-67	Beta
G6PC2	0,653	0,496	0,6268200625	1,1542867051367E-63	Beta
NPTX2	0,571	0,306	1,0415285582	2,61109802635189E-60	Beta
SURF4	0,858	0,822	0,447994563	9,12108033327513E-57	Beta
C9orf3	0,551	0,383	0,3279724129	4,30623830911651E-55	Beta
EIF4A2	0,924	0,86	0,445200542	8,09150272653895E-52	Beta
RRAGD	0,614	0,531	0,3905157769	2,91323673716647E-48	Beta
MSI2	0,524	0,358	0,7339735839	2,17687814634181E-42	Beta
TSPAN13	0,705	0,686	0,4378083583	2,83761197200997E-31	Beta
PPP1R1A	0,845	0,864	0,4260213253	4,19891031343563E-31	Beta
GAPDH	0,996	0,987	0,5180170213	1,72190833957654E-29	Beta
ZNF395	0,602	0,586	0,6222199228	1,99376518871962E-28	Beta

YWHAQ	0,869	0,796	0,2726649564	2,16682660541859E-27	Beta
C1QL1	0,644	0,45	0,6342884224	3,73570329443374E-26	Beta
FBXO16	0,537	0,465	0,7173639099	3,71081272402035E-25	Beta
TSC22D1	0,818	0,692	0,8496934577	1,6422977492076E-23	Beta
VEGFA	0,793	0,745	0,3216421953	4,02446632061277E-20	Beta
GAD2	0,884	0,873	0,6039079339	4,00725841462798E-18	Beta
DYNLT3	0,563	0,513	0,2874510192	4,8254902235567E-15	Beta
SREBF1	0,506	0,466	0,2524269539	3,57919690641169E-13	Beta
INS	0,998	0,962	4,2879155194	1,44178898331504E-12	Beta
DNAJB9	0,666	0,554	0,3666207655	5,06218692978171E-12	Beta
DNAJC12	0,745	0,667	0,295099149	3,62289697131742E-10	Beta
PDX1	0,572	0,517	0,4481838725	7,16067538007134E-09	Beta
IAPP	0,967	0,846	3,6221970091	2,06537970235971E-08	Beta
SYT13	0,573	0,523	0,3604700692	2,32137261958594E-06	Beta
CNP	0,639	0,582	0,2501174903	0,0002347444	Beta
PLCXD3	0,624	0,508	0,260404075	0,0002746961	Beta
HSP90B1	0,866	0,854	0,4559916112	0,0005725487	Beta
SCD	0,717	0,678	0,3806323734	0,0006089205	Beta
RBP4	0,973	0,606	1,4371511168	2,78021606663511E-212	Delta
PCP4	0,828	0,601	0,4691605954	1,3937926210722E-43	Delta
RGS2	0,674	0,288	1,201825629	1,98431797016863E-41	Delta
SEC11C	0,955	0,85	0,6648399306	5,36612693132981E-36	Delta
DHRS2	0,712	0,491	0,450766932	1,38324499582327E-34	Delta
HHEX	0,529	0,327	0,5014926301	1,42340440610336E-29	Delta
SST	0,995	0,885	3,0231334376	2,63640543493176E-12	Delta
PSIP1	0,522	0,358	0,3918935667	2,50662902613556E-09	Delta
NLRP1	0,684	0,696	0,2756040318	0,0002171965	Delta
PMEPA1	0,936	0,516	2,1664103019	2,43980458184273E-295	Ductal
SERPING1	0,848	0,302	2,0441181697	2,6328826954233E-267	Ductal
MMP7	0,756	0,126	2,2650118164	5,86596636333835E-256	Ductal
S100A14	0,735	0,098	1,3914616732	6,9637819926158E-247	Ductal
ATP1A1	0,955	0,788	1,2537392904	7,27942086491743E-245	Ductal
ONECUT2	0,576	0,125	1,2529453426	5,72386147114614E-243	Ductal
ANXA2	0,989	0,913	1,3074369515	5,77950333562933E-238	Ductal
FLNA	0,832	0,277	1,6578837475	3,20057988605874E-236	Ductal
CFTR	0,684	0,104	2,0837802873	1,67429032696877E-230	Ductal
TFPI2	0,626	0,042	1,572984309	1,41027618268196E-223	Ductal
ALDH1A3	0,629	0,084	1,627508757	4,97347070083383E-210	Ductal
PPAP2C	0,693	0,148	1,1571621504	2,2915295801522E-208	Ductal
TPM4	0,874	0,598	0,881820932	8,13948108700579E-200	Ductal
KRT7	0,933	0,672	1,4151265722	6,67706529570305E-199	Ductal
IFITM2	0,824	0,454	1,3059716394	7,16175215514656E-197	Ductal
QSOX1	0,752	0,373	0,8145397581	7,19826204544392E-181	Ductal
TACSTD2	0,952	0,732	1,1467619027	1,59107654899038E-178	Ductal
TPM1	0,913	0,639	1,1077910751	9,31863073029105E-176	Ductal
COL18A1	0,731	0,088	1,3546869518	2,70239099307185E-173	Ductal
ANXA4	0,847	0,435	1,4872271453	3,66837385614964E-171	Ductal
GYPC	0,5	0,022	0,7976943522	4,06278598247728E-171	Ductal
SERPINA5	0,648	0,202	1,7298413871	1,8106146809433E-162	Ductal
DUSP1	0,834	0,543	1,0428537794	5,51587067004299E-161	Ductal
JUP	0,894	0,709	0,9907151882	7,0187546723551E-160	Ductal
ANXA3	0,615	0,151	1,0092449632	1,06488467630793E-158	Ductal
STOM	0,526	0,156	0,779686001	1,08784572301028E-158	Ductal

SERINC2	0,877	0,591	1,2406064003	5,24175424233677E-158	Ductal
LAMC2	0,723	0,305	1,5349806536	2,46501589971136E-155	Ductal
ITGA5	0,56	0,101	0,9759217359	1,34436288007517E-152	Ductal
IER3	0,93	0,647	0,9583412062	1,14242542312951E-151	Ductal
MALL	0,589	0,115	1,1296536897	1,1354898920069E-150	Ductal
WWTR1	0,585	0,159	0,897782743	2,2149688038377E-147	Ductal
SPP1	0,63	0,254	1,4306337148	4,38789908163912E-147	Ductal
ANXA11	0,908	0,827	0,7896263536	1,64884049307359E-143	Ductal
ACTN4	0,962	0,872	0,9669487616	2,99509640066141E-142	Ductal
HSPG2	0,505	0,076	1,0109439907	5,96618885123092E-141	Ductal
CCL2	0,579	0,13	1,7667337028	4,91203143280828E-139	Ductal
ACTB	0,995	0,992	0,8663317718	4,17515470092534E-137	Ductal
TNFAIP2	0,543	0,115	1,3421466142	2,46472190109172E-136	Ductal
SLC2A1	0,554	0,197	0,7832620328	1,30950123066384E-134	Ductal
TMSB10	0,982	0,937	0,9709720055	5,35818925637628E-133	Ductal
CLDN1	0,779	0,518	1,0347263355	1,0508794016132E-132	Ductal
S100A16	0,81	0,518	1,1817936384	4,38552822464946E-129	Ductal
AHNAK	0,644	0,259	1,0369251616	4,22614891591477E-128	Ductal
CTSH	0,633	0,259	1,3328031769	4,06402907521807E-127	Ductal
SDC1	0,529	0,128	0,9045348147	1,34122984174739E-125	Ductal
CCND1	0,559	0,224	0,8640046196	3,0489856586721E-125	Ductal
ELF3	0,93	0,812	0,906350317	1,41416459342134E-120	Ductal
WDR1	0,795	0,606	0,7707634857	1,43380538150709E-120	Ductal
PFN1	0,919	0,85	0,6083836737	8,88199911889081E-118	Ductal
MYOF	0,539	0,151	0,755091394	1,21734237171213E-116	Ductal
S100A11	0,988	0,976	0,7391934456	1,80720759632237E-116	Ductal
TGFA	0,564	0,19	0,8335050703	6,49063795495022E-116	Ductal
ITGA3	0,713	0,399	0,9090411173	2,23889472300785E-114	Ductal
CAPN1	0,8	0,595	0,7997504834	3,66405663136225E-113	Ductal
ITGB5	0,605	0,304	0,5909370894	3,62118472292275E-112	Ductal
TSPAN15	0,656	0,322	0,915575774	4,84118992835638E-111	Ductal
LAD1	0,785	0,618	0,8338035189	5,8492417136323E-111	Ductal
PRDX1	0,899	0,812	0,7736279588	1,44162307309972E-110	Ductal
CITED4	0,628	0,286	0,9137259497	1,91847664891281E-108	Ductal
CNDP2	0,773	0,528	0,5882255565	1,01925730292947E-107	Ductal
DUSP5	0,633	0,303	0,8244062616	9,36815004824065E-107	Ductal
RBPMS	0,781	0,557	0,7608610116	3,27285364397504E-106	Ductal
UGCG	0,641	0,301	0,9190627387	1,59713109182636E-104	Ductal
SMAD3	0,728	0,458	0,7973177163	1,93783723191269E-104	Ductal
TNFRSF12A	0,904	0,849	0,873058682	1,5875615926745E-101	Ductal
CFL1	0,967	0,957	0,4416226634	3,07415104291112E-101	Ductal
PFKFB3	0,73	0,39	1,1452412209	6,88539558793346E-101	Ductal
PPP1R15A	0,716	0,379	0,6360290725	3,94865148232797E-100	Ductal
PDLIM3	0,533	0,174	1,1248017266	4,37202189063817E-100	Ductal
MUC20	0,577	0,22	0,9438433565	3,50923745447757E-99	Ductal
CAV2	0,626	0,294	0,8524588784	2,37108693737584E-98	Ductal
GPRC5B	0,629	0,295	0,942199941	8,03252525954192E-98	Ductal
CEBPB	0,685	0,391	0,9846165452	9,26537097170576E-98	Ductal
NCEH1	0,541	0,212	0,8037817166	2,22357944170224E-97	Ductal
NCK2	0,6	0,298	0,6868639255	3,21869896851177E-97	Ductal
B4GALT5	0,663	0,378	0,8110913375	1,41355336925099E-95	Ductal
CLDN10	0,6	0,331	1,0926759769	1,81418129198472E-93	Ductal
HSPB8	0,575	0,234	0,7726087815	1,22041423644394E-92	Ductal

SLC44A2	0,606	0,305	0,7755375858	1,23216739714832E-92	Ductal
S100A10	0,926	0,586	1,2174160379	8,81520127258491E-92	Ductal
SLC25A23	0,564	0,286	0,493770927	2,338293716899E-91	Ductal
CNN3	0,663	0,388	0,7917884067	3,06133008423182E-89	Ductal
CAPN2	0,809	0,706	0,6545373473	5,69230966855944E-89	Ductal
MGAT4B	0,804	0,658	0,7287524237	1,37762227299486E-84	Ductal
NOTCH2	0,527	0,225	0,6448742698	1,58386248117041E-84	Ductal
CYR61	0,657	0,397	1,348093622	7,89453631758726E-84	Ductal
BTG1	0,902	0,832	0,7995133162	2,02422348276207E-83	Ductal
ARPC1B	0,842	0,684	0,4702400372	3,21217682807318E-81	Ductal
HLA.B	0,685	0,564	0,9886028485	7,20026175946113E-81	Ductal
ENO1	0,973	0,934	0,4723021414	2,70334263614886E-79	Ductal
APLP2	0,911	0,813	0,6414136388	1,24588073265633E-78	Ductal
LCN2	0,816	0,611	1,0811980986	8,51993435845139E-78	Ductal
PLAUR	0,571	0,278	1,0062327307	2,72579289586512E-77	Ductal
NDUFA4	0,828	0,799	0,6217270324	8,67340010217975E-76	Ductal
PDLIM1	0,726	0,56	0,623360514	4,71547237763415E-75	Ductal
TSPO	0,805	0,681	0,6586083737	1,09547893045406E-74	Ductal
SDHA	0,623	0,41	0,4660362206	2,10379755639266E-74	Ductal
DSG2	0,69	0,476	0,6507949302	8,9222173986896E-74	Ductal
ARL6IP1	0,792	0,602	0,5025874178	9,59544355643903E-74	Ductal
EDN1	0,56	0,247	0,9302945775	4,59717110156764E-73	Ductal
PTP4A2	0,849	0,675	0,5653434192	5,30849670548344E-72	Ductal
PRSS8	0,74	0,546	0,8302565113	9,05973827073409E-72	Ductal
PHLDA2	0,544	0,266	0,9334495782	2,78009407103779E-70	Ductal
SLC4A4	0,733	0,293	1,2570482434	1,69692580900192E-69	Ductal
GOLM1	0,717	0,531	0,7121715543	4,62254737213034E-69	Ductal
C1orf106	0,549	0,262	0,6554207102	4,65527612515423E-69	Ductal
CDC42EP1	0,849	0,792	0,677665522	4,10763642154724E-68	Ductal
LGALS3	0,773	0,613	0,6694597244	7,29868016334283E-68	Ductal
KRT18	0,952	0,957	0,6905683689	3,13660386167033E-67	Ductal
FILIP1L	0,507	0,243	0,8565375353	1,98069747696168E-66	Ductal
TNS1	0,522	0,255	0,4162921319	5,81805013255638E-66	Ductal
KRT8	0,956	0,957	0,6982358862	5,85287423844325E-66	Ductal
MYL12B	0,936	0,894	0,4491851042	2,64067192462828E-65	Ductal
SOX4	0,903	0,8	0,6747192823	4,15439961610102E-65	Ductal
BACE2	0,812	0,667	0,5578754218	4,26476467036924E-64	Ductal
KIAA1522	0,776	0,673	0,5575302777	4,13538337267479E-63	Ductal
LAMC1	0,619	0,422	0,6846297625	1,65633830183043E-62	Ductal
AGRN	0,683	0,194	1,0157986737	2,97920354923967E-62	Ductal
LBH	0,546	0,317	0,5303384556	5,78880905039959E-61	Ductal
RHOC	0,905	0,865	0,561454881	3,64842569043459E-58	Ductal
TJP2	0,635	0,415	0,4071544978	6,19855966542591E-58	Ductal
PPIC	0,609	0,426	0,5626879681	2,90321286094061E-56	Ductal
C3	0,684	0,397	0,7105506791	3,48239356363493E-56	Ductal
MYH9	0,9	0,845	0,5898171449	1,45289812330071E-55	Ductal
NAMPT	0,768	0,594	0,8154950774	2,26616953308117E-55	Ductal
ZYX	0,72	0,59	0,7262064609	1,00061132204868E-53	Ductal
TES	0,686	0,55	0,5876734542	1,01421229207319E-53	Ductal
TNIP1	0,674	0,502	0,6617437521	1,01713152231668E-53	Ductal
CAST	0,757	0,664	0,5300106467	1,07617996074763E-53	Ductal
SDCBP	0,731	0,537	0,420208184	1,1040983978991E-53	Ductal
CLDN4	0,903	0,836	0,6781067809	2,06100574930682E-53	Ductal

CREB5	0,5	0,247	0,6408436522	2,51210161534387E-53	Ductal
SYNGR2	0,854	0,824	0,5674270028	2,74459689279213E-53	Ductal
YWHAZ	0,828	0,712	0,5417294119	2,24554023950664E-52	Ductal
TAPBP	0,87	0,757	0,3841023121	3,97007898382776E-52	Ductal
MYL12A	0,872	0,796	0,4729008533	5,25188237478071E-52	Ductal
CXADR	0,621	0,438	0,5935100117	1,71587119928223E-51	Ductal
VCL	0,602	0,416	0,6121979332	3,30240886290562E-50	Ductal
PERP	0,794	0,676	0,4521266436	1,55944586466206E-49	Ductal
CLDN3	0,595	0,412	0,5884367964	2,19101020000355E-49	Ductal
MYO1C	0,744	0,669	0,543501432	6,82552788489556E-49	Ductal
SET	0,866	0,819	0,5091233458	1,01385508131756E-48	Ductal
MYADM	0,693	0,49	0,4227070289	2,61908678215102E-48	Ductal
GNG5	0,863	0,845	0,3971389504	6,93159995384131E-48	Ductal
CTNND1	0,642	0,478	0,7312511	1,9728990741812E-47	Ductal
CTSD	0,902	0,729	1,5964008593	2,6531406933219E-47	Ductal
CBX3	0,723	0,604	0,4895871865	1,1338562521644E-46	Ductal
GIPC1	0,541	0,328	0,3433506762	6,1772192469504E-46	Ductal
MAL2	0,763	0,53	0,6240034679	1,14358027441884E-45	Ductal
EXT1	0,577	0,386	0,546229624	1,34084914261827E-45	Ductal
DYNLT1	0,734	0,628	0,3547022574	3,21724071088042E-45	Ductal
ERGIC1	0,746	0,667	0,6134456405	3,89874104786189E-45	Ductal
LMNA	0,854	0,657	0,27441642	1,08446539816344E-44	Ductal
HDGF	0,881	0,867	0,5462327564	2,39333798546649E-44	Ductal
ATP5B	0,867	0,805	0,4717089282	3,07984287456922E-44	Ductal
ANXA5	0,91	0,726	0,653339057	1,35339440670213E-43	Ductal
ZC3H12A	0,571	0,362	0,6877582731	1,49613562911361E-43	Ductal
ABCC3	0,608	0,44	0,5566994428	2,89105243491028E-43	Ductal
HIF1A	0,74	0,666	0,7082436171	2,96431593269318E-43	Ductal
PRSS22	0,553	0,377	0,6966065989	1,28064503286447E-42	Ductal
DHX15	0,566	0,386	0,3408278257	2,16131489604639E-42	Ductal
KLF6	0,866	0,775	0,574937144	2,90794397143921E-42	Ductal
SEPT9	0,76	0,366	0,5883406617	4,23409135535375E-42	Ductal
GSTP1	0,944	0,944	0,5224362306	4,34974392136407E-42	Ductal
CLDN7	0,915	0,917	0,42146044	1,05639144008496E-41	Ductal
SLC39A1	0,772	0,608	0,2994012944	1,70012553253031E-41	Ductal
TNFSF10	0,605	0,462	0,639875057	3,33316217561739E-41	Ductal
TSPAN3	0,898	0,844	0,3614595432	1,43893529289298E-40	Ductal
RAB10	0,75	0,619	0,351189004	1,50584052025976E-40	Ductal
TPGS2	0,522	0,349	0,504240713	1,155742868837E-39	Ductal
C1orf198	0,516	0,345	0,4835632718	6,85449459544614E-39	Ductal
CTNNA1	0,831	0,613	0,5833718455	7,06950092938082E-39	Ductal
ATP13A3	0,656	0,533	0,6276354557	1,76569282554634E-38	Ductal
TPI1	0,924	0,89	0,2848113004	3,52196339760808E-38	Ductal
RALA	0,526	0,353	0,3193787248	4,29115576564496E-38	Ductal
PTBP1	0,703	0,559	0,3368061101	9,55013190039407E-38	Ductal
CLINT1	0,589	0,406	0,3250693519	2,3947092933593E-37	Ductal
HN1	0,736	0,62	0,4567319401	7,66941644589307E-37	Ductal
PLSCR1	0,619	0,502	0,532829793	1,71437922963705E-36	Ductal
MAP4	0,652	0,552	0,4543551018	2,30774137797975E-36	Ductal
BCL10	0,523	0,358	0,5041991637	8,22871575185186E-36	Ductal
RAP2B	0,514	0,34	0,3642089813	9,94879775615409E-36	Ductal
WWC1	0,533	0,36	0,3367594277	1,14167230539995E-35	Ductal
STK24	0,697	0,553	0,4499697001	3,60661027986938E-35	Ductal

DCDC2	0,516	0,167	0,6812029869	4,31105293108497E-35	Ductal
LAMB3	0,715	0,579	0,5538215684	1,01827216067042E-34	Ductal
PTPRK	0,556	0,431	0,4657889566	2,77313683590291E-34	Ductal
SH3BGRL3	0,876	0,87	0,582783011	4,5524477689416E-34	Ductal
TMEM123	0,734	0,634	0,394016746	1,191359940911E-33	Ductal
MAP1LC3B	0,736	0,703	0,4167239377	4,7759304475817E-33	Ductal
SYPL1	0,592	0,483	0,4249283078	5,76130349834149E-33	Ductal
HN1L	0,618	0,451	0,3278205898	7,68568889455006E-33	Ductal
MYH14	0,636	0,441	0,3315700489	2,68285897071425E-32	Ductal
BCL2L1	0,625	0,46	0,4384603174	2,84757731700997E-32	Ductal
ALDH3A2	0,614	0,495	0,5809693006	3,76517220464758E-32	Ductal
TAX1BP3	0,542	0,388	0,6095555251	4,53784794519289E-32	Ductal
HNRNPA0	0,726	0,569	0,2903420931	4,18404148746017E-31	Ductal
CAPNS1	0,896	0,794	0,3066658547	9,42330796663802E-31	Ductal
SEMA4B	0,619	0,41	0,3235889718	1,50967389073556E-30	Ductal
PPP1CB	0,766	0,656	0,3183669374	1,8932896256781E-30	Ductal
CDKN1A	0,792	0,627	0,3588848742	6,29454014195123E-30	Ductal
PSMA7	0,857	0,872	0,3476896521	1,69181630764049E-29	Ductal
TJP1	0,572	0,456	0,442629568	2,13653222875392E-29	Ductal
TNKS1BP1	0,516	0,387	0,4435929026	2,42503824705265E-29	Ductal
ADAM9	0,648	0,523	0,4386341242	2,61298696019212E-29	Ductal
FDFT1	0,708	0,639	0,5777152067	3,09842698455762E-29	Ductal
HADHB	0,634	0,527	0,3345673825	5,44720370944768E-29	Ductal
PGK1	0,834	0,724	0,2999876351	5,5497901168251E-29	Ductal
F11R	0,643	0,564	0,4341434832	7,2213221163562E-29	Ductal
UBE2Z	0,529	0,364	0,2785758982	1,5689248255644E-28	Ductal
STAT6	0,568	0,453	0,4287823069	3,24421042885715E-28	Ductal
MPRIP	0,517	0,371	0,4404627597	3,59349379433118E-28	Ductal
GALNT1	0,526	0,39	0,4650264207	4,2512235648679E-28	Ductal
CMPK1	0,621	0,519	0,4040015793	9,78370725852332E-28	Ductal
PSMB8	0,542	0,413	0,4261236995	1,28876382181605E-27	Ductal
IGFBP7	0,769	0,794	0,7615795846	1,29206833710666E-27	Ductal
TPM3	0,818	0,709	0,2701034144	1,6788335323103E-27	Ductal
DHCR24	0,523	0,391	0,5376608742	2,8631170950609E-27	Ductal
MSMO1	0,562	0,435	0,4507512077	8,88850622378609E-27	Ductal
PNRC1	0,799	0,757	0,4761905958	1,67021420072784E-26	Ductal
UBALD2	0,674	0,595	0,5395214026	2,15944682448416E-26	Ductal
MACF1	0,659	0,527	0,3783375131	2,92801943343806E-26	Ductal
NECAP2	0,508	0,383	0,4054134056	1,06423438764426E-25	Ductal
HMGA1	0,728	0,654	0,8909425727	2,72392302530499E-25	Ductal
TRIP6	0,522	0,413	0,4657951062	1,020579606668E-24	Ductal
RNPEP	0,549	0,455	0,4362766663	1,06294823367227E-24	Ductal
ARL6IP5	0,65	0,502	0,2798918395	1,24825501127312E-24	Ductal
PIM3	0,717	0,609	0,4976628947	2,63398198719688E-24	Ductal
IL4R	0,568	0,467	0,497062984	3,41430556736633E-24	Ductal
SDC4	0,884	0,833	0,4318823531	1,96696145861189E-23	Ductal
ATF3	0,625	0,548	0,7903605747	3,43848331240079E-23	Ductal
UBE2I	0,739	0,722	0,3678182603	1,8134304782738E-22	Ductal
SEL1L3	0,596	0,487	0,4562028483	2,21947567646481E-22	Ductal
B4GALT1	0,671	0,598	0,4191777516	1,05960412303267E-21	Ductal
FARP1	0,554	0,456	0,3719955837	1,45217343226879E-21	Ductal
PPDPF	0,884	0,816	0,3166923107	2,26678989238009E-21	Ductal
MSN	0,676	0,635	0,5010941191	7,1214227628297E-21	Ductal

VASP	0,561	0,448	0,4595473491	7,79821953711585E-21	Ductal
HNRNPF	0,769	0,769	0,3407851711	1,02311026637009E-20	Ductal
MPZL1	0,615	0,548	0,3650132182	1,30993691942579E-20	Ductal
PFKP	0,725	0,47	0,4258211038	1,39526906919474E-20	Ductal
TUBB	0,912	0,952	0,3879155894	2,09377423899752E-20	Ductal
FOXO3	0,625	0,363	0,4369941227	3,88011444742262E-20	Ductal
COX5B	0,793	0,719	0,2728331518	2,77796700459711E-19	Ductal
EFNA1	0,636	0,548	0,4912783548	4,63193237294288E-19	Ductal
PKP3	0,5	0,385	0,4107798982	5,01185982569046E-19	Ductal
TRIM47	0,519	0,408	0,3932902502	7,75262892853933E-19	Ductal
ATP1B1	0,729	0,636	0,4440630351	1,01725373505798E-18	Ductal
CCDC6	0,538	0,44	0,3687891777	1,01742830149575E-18	Ductal
CERS2	0,732	0,71	0,2943898247	1,30478862108734E-18	Ductal
CDC37	0,69	0,662	0,4070019772	2,12441981525203E-18	Ductal
CTTN	0,614	0,551	0,3600743733	2,98937527795908E-18	Ductal
C8orf4	0,616	0,557	0,7471753432	3,15285273800103E-18	Ductal
IER2	0,834	0,567	0,3191039519	6,61765219869083E-18	Ductal
FAM129B	0,795	0,813	0,4414926436	7,84236987860752E-18	Ductal
CHP1	0,68	0,656	0,3483877176	1,34659510398657E-17	Ductal
TNFRSF10B	0,504	0,411	0,3844921618	2,07171352608555E-17	Ductal
RCC2	0,508	0,413	0,3740687471	2,98259165860871E-17	Ductal
TST	0,511	0,427	0,4032297629	3,69248814805919E-17	Ductal
DHRS3	0,657	0,653	0,6230707225	3,73697133521515E-17	Ductal
TUBB4B	0,792	0,816	0,4946405955	4,70627347738601E-17	Ductal
LAPTM4B	0,777	0,596	0,4338274556	4,86679790795414E-17	Ductal
LMO7	0,548	0,315	0,5490407828	5,05901348434231E-17	Ductal
ARSD	0,524	0,426	0,4068158971	5,08286364637866E-17	Ductal
MYO6	0,578	0,448	0,3488718436	6,91791722478379E-17	Ductal
IQGAP1	0,734	0,547	0,4343609348	1,2217292425177E-16	Ductal
BIRC3	0,569	0,44	0,438228683	2,00277266328545E-16	Ductal
AKAP13	0,598	0,534	0,3635564104	2,65756008990327E-16	Ductal
TMSB4X	0,997	0,997	0,2837039198	2,73774783167942E-16	Ductal
ZFP36L2	0,502	0,407	0,450323578	3,02860762294556E-16	Ductal
ACAA2	0,701	0,489	0,3792175745	7,65614077641924E-16	Ductal
KIFC3	0,526	0,433	0,4280722196	1,84976295864418E-15	Ductal
DDIT4	0,809	0,749	0,5353415868	2,324342054306E-15	Ductal
PXN	0,575	0,51	0,3924640021	5,16823555756214E-15	Ductal
TMEM51	0,627	0,569	0,3775114301	8,8699638011286E-15	Ductal
CLTB	0,597	0,539	0,4132810558	1,26203071142074E-14	Ductal
HPCAL1	0,505	0,417	0,4076891363	1,38526836771765E-14	Ductal
SP100	0,537	0,449	0,2937528822	2,41660583994101E-14	Ductal
CD151	0,707	0,717	0,3881265802	1,02413537972524E-13	Ductal
CTNNB1	0,619	0,441	0,4049978453	1,4008422244134E-13	Ductal
S100A13	0,571	0,497	0,3236148961	2,95048023308581E-13	Ductal
HADHA	0,711	0,741	0,2753689998	4,79478188848273E-13	Ductal
MYO5B	0,529	0,446	0,3030366215	7,39498348024722E-13	Ductal
NFKBIZ	0,645	0,582	0,4128861772	9,07080147248855E-13	Ductal
EIF6	0,65	0,64	0,354188577	9,43108758667128E-13	Ductal
RBM47	0,629	0,568	0,2980497618	9,57141830472188E-13	Ductal
CLMN	0,6	0,589	0,4108690727	2,1376186685878E-12	Ductal
TCP1	0,639	0,646	0,3196621983	2,21908041514642E-12	Ductal
TMEM50A	0,608	0,576	0,2985610998	2,97844554143012E-12	Ductal
GALNT3	0,508	0,407	0,3068875128	3,09369913790235E-12	Ductal

DDAH2	0,608	0,44	0,2624431328	3,67502970181794E-12	Ductal
CSNK1E	0,68	0,475	0,3800365638	4,40273431229699E-12	Ductal
VAMP8	0,839	0,873	0,2539323288	7,3945581124507E-12	Ductal
BHLHE40	0,697	0,653	0,3809683691	1,03998335935122E-11	Ductal
EIF4G1	0,773	0,786	0,3351949152	1,32136084554446E-11	Ductal
NCOA7	0,629	0,487	0,4380408284	2,0872280396262E-11	Ductal
TMBIM1	0,897	0,916	0,2551285464	2,50626441613097E-11	Ductal
STAU1	0,687	0,694	0,2816491523	3,1442868621887E-11	Ductal
GNG12	0,68	0,492	0,2779124253	4,27428551192057E-11	Ductal
XRCC5	0,724	0,628	0,2701459664	5,94252664583236E-11	Ductal
CTBP2	0,552	0,482	0,291295567	6,04789331501011E-11	Ductal
HMGB1	0,869	0,789	0,3020703265	1,38883412772009E-10	Ductal
FLII	0,572	0,54	0,3255515452	2,10837957792595E-10	Ductal
TGIF1	0,543	0,479	0,3308602593	6,01011411866391E-10	Ductal
TMEM87A	0,569	0,512	0,293023822	8,48329527102085E-10	Ductal
DERL1	0,501	0,456	0,3026750922	1,41076164684789E-09	Ductal
ENAH	0,621	0,592	0,36477059	1,60236670415585E-09	Ductal
SUN2	0,572	0,354	0,2831826869	2,35282657016288E-09	Ductal
PPP1R14B	0,728	0,722	0,3819426387	2,63179645357196E-09	Ductal
KRT19	0,953	0,462	2,1585819391	3,64424564925321E-09	Ductal
BRI3	0,854	0,885	0,2725362762	1,10753042627981E-08	Ductal
EIF4G2	0,911	0,823	0,2656868309	1,76392147923269E-08	Ductal
LYPLA2	0,6	0,58	0,3727970873	1,77868645208789E-08	Ductal
JUN	0,867	0,69	0,3342689695	5,57247477556677E-08	Ductal
ADM	0,51	0,414	0,429772353	0,000000066	Ductal
PGD	0,508	0,457	0,3229928877	6,86826438819731E-08	Ductal
NFKBIA	0,775	0,767	0,5383822386	8,03894829565906E-08	Ductal
TFPI	0,515	0,441	0,2889369863	1,7129160067843E-07	Ductal
RELA	0,521	0,493	0,3052408575	1,75389956822097E-07	Ductal
CHMP2B	0,523	0,493	0,2693555261	1,9722138114054E-07	Ductal
HNRNPAB	0,665	0,67	0,293694959	2,23760755415264E-07	Ductal
VDAC2	0,795	0,83	0,2695874506	2,46451246528917E-07	Ductal
PICALM	0,518	0,349	0,2931112229	2,57445519571737E-07	Ductal
LASP1	0,634	0,628	0,2650517495	0,000000278	Ductal
APH1A	0,605	0,609	0,3198259798	3,89715077731062E-07	Ductal
ZFP36L1	0,927	0,932	0,326203611	0,000000542	Ductal
SYNCRIP	0,609	0,585	0,2705976425	5,46721989096314E-07	Ductal
SRSF3	0,723	0,78	0,2890721599	8,05845980736884E-07	Ductal
TINAGL1	0,796	0,038	1,6117658476	8,52873078323075E-07	Ductal
ARHGDIA	0,749	0,789	0,3009962372	9,24574570340368E-07	Ductal
ITGA2	0,588	0,543	0,4233471861	1,11070970174428E-06	Ductal
EIF1AX	0,731	0,761	0,2981089063	1,6758208392043E-06	Ductal
NFE2L2	0,613	0,625	0,2576421997	0,000004921	Ductal
GDE1	0,559	0,568	0,2934452785	4,95877075592781E-06	Ductal
MPST	0,553	0,534	0,3198777132	6,98208770281105E-06	Ductal
ELOVL1	0,629	0,641	0,2963601791	7,59585897170304E-06	Ductal
NFIB	0,573	0,526	0,2605413495	3,25546414639515E-05	Ductal
LITAF	0,872	0,462	1,1030052813	4,70375989093281E-05	Ductal
LAMP2	0,641	0,612	0,2594478851	0,000049262	Ductal
PDLIM5	0,592	0,553	0,2562451342	0,00005558	Ductal
PNPLA2	0,605	0,475	0,2584950065	8,68605719994701E-05	Ductal
BZW1	0,716	0,726	0,2704897055	0,0001870071	Ductal
HEBP1	0,505	0,488	0,2635561605	0,0008399864	Ductal

TM4SF5	0,615	0,032	1,3693346722	3,75712617669899E-54	Epsilon
GHRL	1	0,063	6,540181084	3,99927890161298E-48	Epsilon
ACSL1	1	0,511	2,3696143168	3,7199000616629E-13	Epsilon
SPTSSB	0,962	0,428	1,7635717697	7,76812821368196E-13	Epsilon
NNMT	0,692	0,164	1,8565259231	3,07100461527844E-11	Epsilon
VTN	0,615	0,17	1,5470174046	0,000000067	Epsilon
CDKN2A	0,731	0,211	0,8246919239	5,46083371244829E-06	Epsilon
FRZB	0,731	0,294	1,4454990486	3,35028348011587E-05	Epsilon
PLAC8	0,808	0,318	0,7487469044	7,65374462727078E-05	Epsilon
HEPACAM2	0,808	0,478	1,7692727031	0,0006289325	Epsilon
ACOT7	0,538	0,16	0,8438494236	0,0008850845	Epsilon
SERTM1	0,506	0,087	0,9154259156	8,27520249248414E-135	Gamma
STMN2	0,85	0,63	0,8559542009	5,17330880650638E-59	Gamma
PAX6	0,954	0,835	0,6911687694	1,71777811645179E-51	Gamma
SLC6A4	0,518	0,244	0,7909528048	1,72756456353615E-44	Gamma
ETV1	0,673	0,421	0,693020194	3,48195757560133E-26	Gamma
SCG2	0,973	0,935	0,4596019124	9,94065170371895E-25	Gamma
CARD11	0,53	0,272	0,5274642499	4,51454976285447E-18	Gamma
ID2	0,835	0,693	0,5605926407	3,97816166434255E-17	Gamma
SCGB2A1	0,661	0,499	0,4227061326	2,23261284766201E-16	Gamma
PPY	0,995	0,577	6,1621226941	4,69402205059878E-13	Gamma
AQP3	0,818	0,759	1,1388036203	7,53453354879978E-13	Gamma
C16orf45	0,63	0,498	0,4366386787	2,72840432662841E-12	Gamma
DDR1	0,908	0,842	0,3514192831	9,18256980250636E-12	Gamma
INPP5F	0,571	0,384	0,3575533459	4,19332497138101E-11	Gamma
PTP4A3	0,709	0,577	0,5087284288	1,8234668024323E-10	Gamma
ID4	0,632	0,503	0,5108819437	2,34537968818928E-10	Gamma
MALAT1	0,976	0,97	0,469634952	5,72085783042863E-06	Gamma
FXYD2	0,811	0,72	0,3588950844	0,0004297099	Gamma

Supplemental table 4

Set of genes regulated in *pax6b*$^{-/-}$ mutant in PECs and EECs which also present a close pax6b DNA-binding site (<= 10kb).

Pax6 bound

Gene	Pax6.GRN	Gene	Pax6.GRN	Gene	Pax6.GRN
abcc9	Both	hapln1a	Both	pou3f3a	Both
acmsd	Both	her4.1	Both	PPP1R3G	Both
adcyap1b	Both	hmx4	Both	ptn	Both
aes	Both	hsd17b12a	Both	rnd1b	Both
alcamb	Both	large2	Both	rnd2	Both
arid5b	Both	lenep	Both	rnf19a	Both
atf2	Both	lipg	Both	rorab	Both
atp1a3a	Both	lmo4a	Both	rtn4rl2b	Both
bambib	Both	map2k6	Both	scrt1a	Both
BX284638.1	Both	mdkb	Both	si:ch211-255p10.3	Both
BX530077.1	Both	meis1b	Both	si:dkey-280e21.3	Both
cbfa2t2	Both	nat8l	Both	slc7a7	Both
cd9b	Both	neurod1	Both	sox4a	Both
cdkn1ca	Both	nid2a	Both	sp8a	Both
copz2	Both	nlgn4a	Both	sprn2	Both
ctnnd2b	Both	nog1	Both	st5	Both
dcxr	Both	nrp2b	Both	syt6a	Both
dll4	Both	odc1	Both	tle3a	Both
eif4ebp3l	Both	olfml3a	Both	trdn	Both
elp4	Both	pard3bb	Both	yjefn3	Both
etv5b	Both	pdgfbb	Both		
fam49al	Both	pfkfb3	Both		
flna	Both	plch1	Both		
fstl1a	Both	plcxd3	Both		
gadd45ga	Both	plod2	Both		
ackr3a	PECs	hoxa3a	PECs	robo2	PECs
ackr3b	PECs	hoxc5a	PECs	roraa	PECs
acot11a	PECs	hoxc6a	PECs	rorcb	PECs
adgrv1	PECs	hs3st3b1b	PECs	rps12	PECs
adora2aa	PECs	id1	PECs	rps13	PECs
afdna	PECs	igsf9b	PECs	rps19	PECs
aff4	PECs	il17rd	PECs	rps28	PECs
afg1lb	PECs	il4r.1	PECs	runx1t1	PECs
ak3	PECs	im:7152348	PECs	sall1a	PECs
AMOTL1	PECs	INAVA	PECs	sat1a.2	PECs
ANKFN1	PECs	insm1b	PECs	scarb2a	PECs
apln	PECs	iqgap2	PECs	scrt2	PECs
arl3l1	PECs	irx3b	PECs	sema3gb	PECs
arl4aa	PECs	jarid2a	PECs	sema4ba	PECs
arpp21	PECs	kcnd3	PECs	sema5a	PECs
asb15a	PECs	kctd12b	PECs	sept12	PECs
ascl1a	PECs	kdm7ab	PECs	sept9a	PECs
ascl1b	PECs	kirrel3l	PECs	serf2	PECs
asf1ba	PECs	klf12b	PECs	serpinh1b	PECs
atp10a	PECs	kmt2a	PECs	shisa9b	PECs
atp1b4	PECs	larp1	PECs	shroom3	PECs
azin1b	PECs	lbx1b	PECs	si:ch1073-209e23.	PECs
baz1b	PECs	ldb1b	PECs	si:ch1073-345a8.1	PECs
bcas3	PECs	ldha	PECs	si:ch211-149l1.2	PECs
bmp1a	PECs	lfng	PECs	si:ch211-161c3.5	PECs
bnip3lb	PECs	lhfpl6	PECs	si:ch211-172l8.4	PECs
bod1l1	PECs	lima1a	PECs	si:ch211-197l9.5	PECs

btbd3a	PECs	llgl1	PECs	si:ch211-204c21.1	PECs
BX927244.1	PECs	LO016987.2	PECs	si:ch211-216b21.2	PECs
c1qtnf12	PECs	lrp2a	PECs	si:ch211-219a15.3	PECs
CABZ01073963.1	PECs	lypd6	PECs	si:ch211-242b18.1	PECs
CABZ01079241.1	PECs	lypd6b	PECs	si:ch211-243j20.2	PECs
cd82a	PECs	mafa	PECs	si:ch211-246i5.5	PECs
cdkn1cb	PECs	mafb	PECs	si:ch211-282k23.2	PECs
celf2	PECs	map3k10	PECs	si:ch73-63e15.2	PECs
cep83	PECs	mapk15	PECs	si:dkey-195m11.8	PECs
cers2a	PECs	mapkbp1	PECs	si:dkey-209n16.2	PECs
chrnb1	PECs	marcksl1b	PECs	si:dkey-21c1.8	PECs
chrnb3a	PECs	mid1ip1l	PECs	si:dkey-225f23.5	PECs
cited2	PECs	mknk2b	PECs	si:dkey-256e7.5	PECs
cldn5a	PECs	mtus1b	PECs	si:dkey-261m9.12	PECs
cntfr	PECs	myb	PECs	si:dkey-49n23.1	PECs
coq2	PECs	mych	PECs	si:dkey-7j22.1	PECs
cplx2	PECs	mydgf	PECs	si:dkey-91m11.5	PECs
CR855996.2	PECs	myo10l1	PECs	si:dkeyp-117b11.3	PECs
crabp2a	PECs	myo15b	PECs	si:dkeyp-41f9.4	PECs
crabp2b	PECs	myt1b	PECs	si:rp71-17i16.6	PECs
CRACR2A	PECs	nanos1	PECs	si:rp71-62i8.1	PECs
crb2a	PECs	nap1l1	PECs	six3a	PECs
CT573817.1	PECs	nbn	PECs	si:zfos-223e1.2	PECs
cxcl12b	PECs	ncoa1	PECs	slc22a23	PECs
cyp1c2	PECs	nelfa	PECs	slc25a24	PECs
cyth1b	PECs	nfat5a	PECs	slc30a2	PECs
dapk3	PECs	nfatc3a	PECs	slc39a13	PECs
ddhd1a	PECs	nhlh2	PECs	slc3a2a	PECs
dennd2da	PECs	nkd1	PECs	slc46a2	PECs
dgkzb	PECs	nlrc3l1	PECs	slc6a22.1	PECs
dlc	PECs	notch1a	PECs	slc6a9	PECs
dpydb	PECs	nova2	PECs	slc9a6a	PECs
E2F2	PECs	npb	PECs	smad1	PECs
ebf3a	PECs	npr1a	PECs	sncb	PECs
ednrab	PECs	nr0b2a	PECs	sncgb	PECs
eef2k	PECs	nr5a2	PECs	snx18a	PECs
efnb2a	PECs	nrip1b	PECs	sobpb	PECs
egf	PECs	nrxn3b	PECs	sort1a	PECs
egr3	PECs	oct2	PECs	sox13	PECs
elovl4b	PECs	otpa	PECs	sox1a	PECs
eml1	PECs	otud7b	PECs	sptlc2b	PECs
emx2	PECs	p4ha2	PECs	ssbp2	PECs
enah	PECs	palm1b	PECs	st8sia1	PECs
enc2	PECs	papss2b	PECs	stac3	PECs
epas1b	PECs	pax3a	PECs	stat3	PECs
epha3	PECs	pax6a	PECs	stim1a	PECs
etaa1	PECs	pcdh18b	PECs	ston2	PECs
eva1a	PECs	pcgf1	PECs	stox1	PECs
eya2	PECs	pdgfab	PECs	syne2b	PECs
fabp11a	PECs	PHF21B	PECs	tagln3a	PECs
fam212b	PECs	phf23b	PECs	tanc2a	PECs
fam214b	PECs	pik3c2b	PECs	tbc1d9	PECs
fam228a	PECs	plce1	PECs	tbx2b	PECs

fam69c	PECs	pleca	PECs	tcf12	PECs
farp1	PECs	pltp	PECs	tenm3	PECs
fbxo4	PECs	plxnb1a	PECs	theg	PECs
fgfr1b	PECs	polr2a	PECs	tiparp	PECs
fgfr2	PECs	pou2f2a	PECs	tmem151a	PECs
fgfr4	PECs	pou3f1	PECs	tmem256	PECs
flrt3	PECs	pou4f1	PECs	tmem47	PECs
foxa	PECs	ppardb	PECs	tnrc18	PECs
foxb1b	PECs	ppox	PECs	trmt9b	PECs
foxg1b	PECs	ppp1r14ba	PECs	trpc3	PECs
foxo3b	PECs	ppp2r2ba	PECs	trpm1b	PECs
FQ311908.1	PECs	prdm14	PECs	tshz3a	PECs
fsta	PECs	prdx5	PECs	ttc26	PECs
fstl1b	PECs	prickle1b	PECs	tuba1c	PECs
gabrb1	PECs	prkcbp1l	PECs	tulp1a	PECs
gabrb2	PECs	prmt6	PECs	ucn3l	PECs
gadd45gb.1	PECs	prrg2	PECs	ush1c	PECs
gapdhs	PECs	prtga	PECs	usp45	PECs
gas1a	PECs	ptch2	PECs	usp8	PECs
gas2b	PECs	pth1rb	PECs	ust	PECs
glceb	PECs	ptmaa	PECs	vax1	PECs
gna11a	PECs	ptmab	PECs	vezf1a	PECs
gnai1	PECs	ptpn11b	PECs	vezf1b	PECs
gpam	PECs	ptprt	PECs	vim	PECs
gpm6aa	PECs	qsox2	PECs	vstm2l	PECs
gpr158a	PECs	rab1ab	PECs	wdr19	PECs
grapa	PECs	rad1	PECs	wnt7aa	PECs
gria1a	PECs	ralbp1	PECs	ypel3	PECs
gria3a	PECs	rarab	PECs	zeb1a	PECs
gsx1	PECs	rasgrp3	PECs	zeb1b	PECs
hbegfa	PECs	rassf7a	PECs	zeb2a	PECs
hepacama	PECs	rbpjb	PECs	zgc:100920	PECs
her12	PECs	rcn3	PECs	zgc:110045	PECs
her15.1	PECs	rdh10a	PECs	zgc:153372	PECs
her15.2	PECs	rergla	PECs	zgc:162939	PECs
her2	PECs	rgma	PECs	zic2a	PECs
her3	PECs	rhbdl3	PECs	zic5	PECs
her4.2	PECs	rhoua	PECs	znf703	PECs
hif1an	PECs	rln1	PECs	znrf1	PECs
hmx2	PECs	rnf24	PECs		
homeza	PECs	rnf38	PECs		
abcc5	EECs	fzd9b	EECs	rimkla	EECs
abhd3	EECs	galn	EECs	rlbp1a	EECs
acana	EECs	gas8	EECs	robo2	EECs
acsl5	EECs	gata2a	EECs	rx3	EECs
adamts18	EECs	gclm	EECs	scarb2c	EECs
angpt2b	EECs	gjd1a	EECs	schip1	EECs
angptl4	EECs	gnpat	EECs	sema3fa	EECs
ankrd6b	EECs	gpr78a	EECs	sfxn1	EECs
ANPEP	EECs	hdac4	EECs	si:ch211-137a8.4	EECs
atp1b3a	EECs	hey1	EECs	si:ch211-148f13.1	EECs
atp6v0a2a	EECs	hic2	EECs	si:ch211-195b13.1	EECs
auts2a	EECs	hmga1b	EECs	si:ch211-251b21.1	EECs

avpr2ab	EECs	hoxb6a	EECs	si:ch211-95j8.3	EECs
CABZ01075125.1	EECs	hoxb9a	EECs	si:ch73-386h18.1	EECs
CABZ01087623.1	EECs	insm1a	EECs	si:dkey-30h22.11	EECs
CABZ01117603.1	EECs	ism1	EECs	si:dkey-85n7.6	EECs
cadps2	EECs	itpkca	EECs	slc1a5	EECs
cbx4	EECs	junbb	EECs	slc38a2	EECs
ccdc85ca	EECs	kirrel1a	EECs	slc38a3a	EECs
ccnb1	EECs	kynu	EECs	sncga	EECs
ccne2	EECs	lhfpl5a	EECs	sox1b	EECs
cdc14b	EECs	LO017965.1	EECs	spen	EECs
cers2b	EECs	lrrc38b	EECs	spry2	EECs
chl1b	EECs	magi1a	EECs	tbrg4	EECs
col12a1a	EECs	mki67	EECs	tbx3a	EECs
col9a1b	EECs	mlip	EECs	timm10	EECs
cspg5a	EECs	mllt11	EECs	tmem132e	EECs
csrnp1b	EECs	mplkip	EECs	tmem35	EECs
cthl	EECs	mycn	EECs	tmlhe	EECs
CU694264.1	EECs	nmt2	EECs	tnpo1	EECs
CU929160.1	EECs	npy2rl	EECs	traf4a	EECs
dhx32b	EECs	nr1d1	EECs	trim3b	EECs
dse	EECs	nt5dc2	EECs	ttll6	EECs
efna2b	EECs	nyap2b	EECs	ttyh3b	EECs
entpd2a.1	EECs	oc90	EECs	txnipa	EECs
EPB41L2	EECs	pacsin1b	EECs	tyms	EECs
epb41l4a	EECs	pax2a	EECs	urahb	EECs
epha4a	EECs	pax6b	EECs	wbp1la	EECs
fam13a	EECs	pik3ip1	EECs	wdr76	EECs
fech	EECs	pld2	EECs	ywhaqa	EECs
fgfr3	EECs	plpp1a	EECs	zgc:101699	EECs
flvcr2b	EECs	pqlc2	EECs	zgc:158689	EECs
FP236812.1	EECs	prkd1	EECs	zgc:165555	EECs
fzd10	EECs	proca1	EECs	zmp:0000000760	EECs
fzd7a	EECs	ptprt	EECs		
fzd9a	EECs	rgcc	EECs		

Supplemental table 5

Set of transcription factors regulated in *pax6b*$^{-/-}$ mutant in PECs and EECs which also present a close pax6b DNA-binding site (<= 10kb).

TFs bound

Gene	Pax6.GRN	Gene	Pax6.GRN	Gene	Pax6.GRN
meis1b	Both	tcf12	PECs	prdm14	PECs
rorab	Both	egr3	PECs	scrt2	PECs
scrt1a	Both	zeb1a	PECs	mych	PECs
arid5b	Both	ascl1a	PECs	sox1a	PECs
pou3f3a	Both	her2	PECs	tulp1a	PECs
sox4a	Both	hmx2	PECs	jarid2a	PECs
atf2	Both	nfat5a	PECs	sall1a	PECs
etv5b	Both	nr5a2	PECs	mafb	PECs
hmx4	Both	vezf1a	PECs	insm1b	PECs
sp8a	Both	CABZ01079241.1	PECs	lbx1b	PECs
neurod1	Both	ascl1b	PECs	nhlh2	PECs
pou2f2a	PECs	foxb1b	PECs	id1	PECs
ppardb	PECs	vezf1b	PECs	rarab	PECs
hoxc6a	PECs	zeb2a	PECs	sox13	PECs
E2F2	PECs	nr0b2a	PECs	epas1b	PECs
nfatc3a	PECs	six3a	PECs	rx3	EECs
myt1b	PECs	emx2	PECs	hoxb9a	EECs
vax1	PECs	foxa	PECs	insm1a	EECs
irx3b	PECs	ebf3a	PECs	junbb	EECs
zic5	PECs	homeza	PECs	nr1d1	EECs
foxo3b	PECs	aff4	PECs	pax2a	EECs
foxg1b	PECs	gsx1	PECs	sox1b	EECs
myb	PECs	hoxa3a	PECs	csrnp1b	EECs
zeb1b	PECs	pou3f1	PECs	tbx3a	EECs
hoxc5a	PECs	tbx2b	PECs	gata2a	EECs
pax3a	PECs	her12	PECs	mycn	EECs
smad1	PECs	roraa	PECs	hey1	EECs
otpa	PECs	mafa	PECs	hic2	EECs
rorcb	PECs	stat3	PECs	hoxb6a	EECs
pou4f1	PECs	zic2a	PECs	pax6b	EECs
pax6a	PECs	her3	PECs	hmga1b	EECs
rbpjb	PECs	klf12b	PECs		

Bibliography

1. Hounnou G, Destrieux C, Desmé J, Bertrand P, Velut S. Anatomical study of the length of the human intestine. Surg Radiol Anat. 2002;24:290–4. doi:10.1007/s00276-002-0057-y.

2. MeSH browser. https://meshb.nlm.nih.gov/search. Accessed 3 Sep 2019.

3. Punj A. Secretions of human salivary gland. Salivary Glands - New Approaches in Diagnostics and Treatment. 2018. doi:10.5772/intechopen.75538.

4. Mese H, Matsuo R. Salivary secretion, taste and hyposalivation. Journal of Oral Rehabilitation. 2007;34:711–23. doi:10.1111/j.1365-2842.2007.01794.x.

5. Information NC for B, Pike USNL of M8R, MD B, Usa 2. How does the stomach work? Institute for Quality; Efficiency in Health Care (IQWiG); 2016. https://www.ncbi.nlm.nih.gov/books/NBK279304/. Accessed 2 Sep 2019.

6. Ramsay PT, Carr A. Gastric acid and digestive physiology. Surg Clin North Am. 2011;91:977–82.

7. Hsu M, Lui F. Physiology, stomach. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2019. http://www.ncbi.nlm.nih.gov/books/NBK535425/. Accessed 2 Sep 2019.

8. Pandol SJ. Regulation of whole-organ pancreatic secretion. Morgan & Claypool Life Sciences; 2010. https://www.ncbi.nlm.nih.gov/books/NBK54132/. Accessed 3 Sep 2019.

9. Park J-h, Kotani T, Konno T, Setiawan J, Kitamura Y, Imada S, et al. Promotion of intestinal epithelial cell turnover by commensal bacteria: Role of short-chain fatty acids. PLoS One. 2016;11. doi:10.1371/journal.pone.0156334.

10. Mayhew TM, Myklebust R, Whybrow A, Jenkins R. Epithelial integrity, cell death and cell loss in mammalian small intestine. Histol Histopathol. 1999;14:257–67.

11. Kiela PR, Ghishan FK. Physiology of intestinal absorption and secretion. Best Pract Res Clin Gastroenterol. 2016;30:145–59. doi:10.1016/j.bpg.2016.02.007.

12. Leushacke M, Barker N. Ex vivo culture of the intestinal epithelium: Strategies and applications. Gut. 2014;63:1345–54.

13. Moran BJ, Jackson AA. Function of the human colon. Br J Surg. 1992;79:1132–7.

14. Phillips M, Patel A, Meredith P, Will O, Brassett C. Segmental colonic length and mobility. Ann R Coll Surg Engl. 2015;97:439–44.

15. Roager HM, Hansen LBS, Bahl MI, Frandsen HL, Carvalho V, Gøbel RJ, et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nat Microbiol. 2016;1:16093.

16. Lake JI, Heuckeroth RO. Enteric nervous system development: Migration, differentiation, and disease. Am J Physiol Gastrointest Liver Physiol. 2013;305:G1–G24. doi:10.1152/ajpgi.00452.2012.

17. Furness JB. The enteric nervous system and neurogastroenterology. Nat Rev Gastroenterol Hepatol. 2012;9:286–94.

18. Goyal RK, Hirano I. The enteric nervous system. New England Journal of Medicine. 1996;334:1106–15. doi:10.1056/NEJM199604253341707.

19. Furness JB. Integrated neural and endocrine control of gastrointestinal function. In: Brierley S, Costa M, editors. The enteric nervous system: 30 years later. Cham: Springer International Publishing; 2016. pp. 159–73. doi:10.1007/978-3-319-27592-5_16.

20. Hennig GW. Spatio-temporal mapping and the enteric nervous system. In: Brierley S, Costa M, editors. The enteric nervous system: 30 years later. Cham: Springer International Publishing; 2016. pp. 31–42. doi:10.1007/978-3-319-27592-5_4.

21. Posovszky C. Development and anatomy of the enteroendocrine system in humans. Developmental Biology of Gastrointestinal Hormones. 2017;32:20–37. doi:10.1159/000475729.

22. Rehfeld JF. A centenary of gastrointestinal endocrinology. Horm Metab Res. 2004;36:735–41. doi:10.1055/s-2004-826154.

23. Latorre R, Sternini C, De Giorgio R, Greenwood-Van Meerveld B. Enteroendocrine cells: A review of their role in brain-gut communication. Neurogastroenterol Motil. 2016;28:620–30. doi:10.1111/nmo.12754.

24. Feher J. 8.3 - intestinal and colonic chemoreception and motility. In: Feher J, editor. Quantitative human physiology (second edition). Boston: Academic Press; 2017. pp. 796–809. doi:10.1016/B978-0-12-800883-6.00079-3.

25. Dockray GJ, Varro A, Dimaline R. Gastric endocrine cells: Gene expression, processing, and targeting of active products. Physiol Rev. 1996;76:767–98.

26. Spence JR, Mayhew CN, Rankin SA, Kuhar M, Vallance JE, Tolle K, et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature. 2011;470:105–9. doi:10.1038/nature09691.

27. Shroyer NF, Kocoshis SA. 31 - anatomy and physiology of the small and large intestines. In: Wyllie R, Hyams JS, editors. Pediatric gastrointestinal and liver disease (fourth edition). Saint Louis: W.B. Saunders; 2011. pp. 324–336.e2. doi:10.1016/B978-1-4377-0774-8.10031-4.

28. Habib AM, Richards P, Cairns LS, Rogers GJ, Bannon CAM, Parker HE, et al. Overlap of endocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. Endocrinology. 2012;153:3054–65.

29. Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, et al. A single-cell survey of the small intestinal epithelium. Nature. 2017;551:333–9. doi:10.1038/nature24489.

30. Low JT, Shukla A, Thorn P. Pancreatic acinar cell: New insights into the control of secretion. The International Journal of Biochemistry & Cell Biology. 2010;42:1586–9. doi:10.1016/j.biocel.2010.07.006.

31. Lee MG, Ohana E, Park HW, Yang D, Muallem S. Molecular mechanism of pancreatic and salivary glands fluid and HCO3- secretion. Physiol Rev. 2012;92:39–74. doi:10.1152/physrev.00011.2011.

32. Medical physiology - 3rd edition. https://www.elsevier.com/books/medical-physiology/boron/978-1-4557-4377-3. Accessed 11 Sep 2019.

33. Jamieson JD, Palade GE. INTRACELLULAR TRANSPORT OF SECRETORY PROTEINS IN THE PANCREATIC EXOCRINE CELL : II. Transport to condensing vacuoles and zymogen granules. The Journal of Cell Biology. 1967;34:597. https://www.ncbi.nlm.nih.g

ov/pmc/articles/PMC2107311/. Accessed 12 Sep 2019.

34. Wäsle B, Edwardson JM. The regulation of exocytosis in the pancreatic acinar cell. Cellular Signalling. 2002;14:191–7. doi:10.1016/S0898-6568(01)00257-1.

35. Jamieson JD, Palade GE. SYNTHESIS, INTRACELLULAR TRANSPORT, AND DISCHARGE OF SECRETORY PROTEINS IN STIMULATED PANCREATIC EXOCRINE CELLS. J Cell Biol. 1971;50:135–58. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2108418/. Accessed 11 Sep 2019.

36. Pandol SJ. The exocrine pancreas. San Rafael (CA): Morgan & Claypool Life Sciences; 2010. http://www.ncbi.nlm.nih.gov/books/NBK54128/. Accessed 12 Sep 2019.

37. Ishiguro H, Naruse S, Steward MC, Kitagawa M, Ko SBH, Hayakawa T, et al. Fluid secretion in interlobular ducts isolated from guinea-pig pancreas. J Physiol. 1998;511 Pt 2:407–22. doi:10.1111/j.1469-7793.1998.407bh.x.

38. Sohma Y, Gray MA, Imai Y, Argent BE. HCO3- transport in a mathematical model of the pancreatic ductal epithelium. J Membr Biol. 2000;176:77–100.

39. Park HW, Lee MG. Transepithelial bicarbonate secretion: Lessons from the pancreas. Cold Spring Harb Perspect Med. 2012;2. doi:10.1101/cshperspect.a009571.

40. Afroze S, Meng F, Jensen K, McDaniel K, Rahal K, Onori P, et al. The physiological roles of secretin and its receptor. Ann Transl Med. 2013;1:29.

41. Röder PV, Wu B, Liu Y, Han W. Pancreatic regulation of glucose homeostasis. Exp Mol Med. 2016;48:e219. doi:10.1038/emm.2016.6.

42. Meier JJ. Chapter 32 - insulin secretion. In: Jameson JL, De Groot LJ, Kretser DM de, Giudice LC, Grossman AB, Melmed S, et al., editors. Endocrinology: Adult and pediatric (seventh edition). Philadelphia: W.B. Saunders; 2016. pp. 546–555.e5. doi:10.1016/B978-0-323-18907-1.00032-9.

43. Jansson L, Barbu A, Bodin B, Drott CJ, Espes D, Gao X, et al. Pancreatic islet blood flow and its measurement. Ups J Med Sci. 2016;121:81–95. doi:10.3109/03009734.2016.1164769.

44. Da Silva Xavier G. The cells of the islets of langerhans. J Clin Med. 2018;7.

doi:10.3390/jcm7030054.

45. Steiner DJ, Kim A, Miller K, Hara M. Pancreatic islet plasticity: Interspecies comparison of islet architecture and composition. Islets. 2010;2:135–45. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2908252/. Accessed 13 Sep 2019.

46. Bloom SR, Polak JM. Somatostatin. Br Med J (Clin Res Ed). 1987;295:288–90. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1247137/. Accessed 16 Sep 2019.

47. Huang X-Q. Somatostatin: Likely the most widely effective gastrointestinal hormone in the human body. World J Gastroenterol. 1997;3:201–4. doi:10.3748/wjg.v3.i4.201.

48. Krejs GJ. Physiological role of somatostatin in the digestive tract: Gastric acid secretion, intestinal absorption, and motility. Scand J Gastroenterol Suppl. 1986;119:47–53.

49. O'Toole TJ, Sharma S. Physiology, somatostatin. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2019. http://www.ncbi.nlm.nih.gov/books/NBK538327/. Accessed 16 Sep 2019.

50. Lonovics J, Devitt P, Watson LC, Rayford PL, Thompson JC. Pancreatic polypeptide: A review. Arch Surg. 1981;116:1256–64. doi:10.1001/archsurg.1981.01380220010002.

51. Vinik A, Feliberti E, Perry RR. Pancreatic polypeptide (PPoma). In: Feingold KR, Anawalt B, Boyce A, Chrousos G, Dungan K, Grossman A, et al., editors. Endotext. South Dartmouth (MA): MDText.com, Inc.; 2000. http://www.ncbi.nlm.nih.gov/books/NBK279067/. Accessed 16 Sep 2019.

52. Scott R, Tan T, Bloom S. Gut hormones and obesity: Physiology and therapies. Vitam Horm. 2013;91:143–94.

53. Pradhan G, Samson SL, Sun Y. Ghrelin: Much more than a hunger hormone. Curr Opin Clin Nutr Metab Care. 2013;16:619–24. doi:10.1097/MCO.0b013e328365b9be.

54. McMillin JM. Blood glucose. In: Walker HK, Hall WD, Hurst JW, editors. Clinical methods: The history, physical, and laboratory examinations. 3rd edition. Boston: Butterworths; 1990. http://www.ncbi.nlm.nih.gov/books/NBK248/. Accessed 16 Sep 2019.

55. McCulloch LJ, Bunt M van de, Braun M, Frayn KN, Clark A, Gloyn AL. GLUT2

(SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: Implications for understanding genetic association signals at this locus. Mol Genet Metab. 2011;104:648–53.

56. Rutter GA, Pullen TJ, Hodson DJ, Martinez-Sanchez A. Pancreatic β-cell identity, glucose sensing and the control of insulin secretion. Biochem J. 2015;466:203–18.

57. Ashcroft FM. ATP-sensitive potassium channelopathies: Focus on insulin secretion. J Clin Invest. 2005;115:2047–58. doi:10.1172/JCI25495.

58. Hou JC, Min L, Pessin JE. Insulin granule biogenesis, trafficking and exocytosis. Vitam Horm. 2009;80:473–506.

59. Han Y-E, Chun JN, Kwon MJ, Ji Y-S, Jeong M-H, Kim H-H, et al. Endocytosis of KATP channels drives glucose-stimulated excitation of pancreatic β cells. Cell Rep. 2018;22:471–81.

60. Quesada I, Tudurí E, Ripoll C, Nadal A. Physiology of the pancreatic alpha-cell and glucagon secretion: Role in glucose homeostasis and diabetes. J Endocrinol. 2008;199:5–19.

61. Baron AD, Brechtel G, Wallace P, Edelman SV. Rates and tissue sites of non-insulin- and insulin-mediated glucose uptake in humans. American Journal of Physiology-Endocrinology and Metabolism. 1988;255:E769–74. doi:10.1152/ajpendo.1988.255.6.E769.

62. Navale AM, Paranjape AN. Glucose transporters: Physiological and pathological roles. Biophys Rev. 2016;8:5–9. doi:10.1007/s12551-015-0186-2.

63. Vargas E, Podder V, Carrillo Sepulveda MA. Physiology, glucose transporter type 4 (GLUT4). In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2019. http://www.ncbi.nlm.nih.gov/books/NBK537322/. Accessed 16 Sep 2019.

64. Sekine N, Cirulli V, Regazzi R, Brown LJ, Gine E, Tamarit-Rodriguez J, et al. Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing. J Biol Chem. 1994;269:4895–902.

65. Quintens R, Hendrickx N, Lemaire K, Schuit F. Why expression of some genes is disallowed in beta-cells. Biochem Soc Trans. 2008;36 Pt 3:300–5.

66. Pullen TJ, Khan AM, Barton G, Butcher SA, Sun G, Rutter GA. Identification of genes selectively disallowed in the pancreatic islet. Islets. 2010;2:89–95.

67. Lemaire K, Granvik M, Schraenen A, Goyvaerts L, Van Lommel L, Gómez-Ruiz A, et al. How stable is repression of disallowed genes in pancreatic islets in response to metabolic stress? PLoS One. 2017;12. doi:10.1371/journal.pone.0181651.

68. Rutter GA. Controlling the identity of the adult pancreatic β cell. Nat Rev Endocrinol. 2017;13:129–30. doi:10.1038/nrendo.2017.1.

69. Pullen TJ, Rutter GA. When less is more: The forbidden fruits of gene repression in the adult β-cell. Diabetes, Obesity and Metabolism. 2013;15:503–12. doi:10.1111/dom.12029.

70. Pullen TJ, Huising MO, Rutter GA. Analysis of purified pancreatic islet beta and alpha cell transcriptomes reveals 11β-hydroxysteroid dehydrogenase (hsd11b1) as a novel disallowed gene. Front Genet. 2017;8. doi:10.3389/fgene.2017.00041.

71. Kalra S, Zargar AH, Jain SM, Sethi B, Chowdhury S, Singh AK, et al. Diabetes insipidus: The other diabetes. Indian J Endocrinol Metab. 2016;20:9–21. doi:10.4103/2230-8210.172273.

72. Halban PA, Polonsky KS, Bowden DW, Hawkins MA, Ling C, Mather KJ, et al. Beta-cell failure in type 2 diabetes: Postulated mechanisms and prospects for prevention and treatment. Diabetes Care. 2014;37:1751–8. doi:10.2337/dc14-0396.

73. Kahn SE. The importance of β-cell failure in the development and progression of type 2 diabetes. J Clin Endocrinol Metab. 2001;86:4047–58. doi:10.1210/jcem.86.9.7713.

74. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2010;33 Suppl 1:S62–9. doi:10.2337/dc10-S062.

75. Kharroubi AT, Darwish HM. Diabetes mellitus: The epidemic of the century. World J Diabetes. 2015;6:850–67. doi:10.4239/wjd.v6.i6.850.

76. Meyers JR. Zebrafish: Development of a vertebrate model organism. Current Protocols Essential Laboratory Techniques. 2018;16:e19. doi:10.1002/cpet.19.

77. Haffter P, Odenthal J, Mullins MC, Lin S, Farrell MJ, Vogelsang E, et al. Mutations

affecting pigmentation and shape of the adult zebrafish. Dev Genes Evol. 1996;206:260–76.

78. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013;496:498–503.

79. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39:e82.

80. Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, et al. Heritable and precise zebrafish genome editing using a CRISPR-cas system. PLoS ONE. 2013;8:e68708.

81. Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, et al. Resolution of ray-finned fish phylogeny and timing of diversification. Proc Natl Acad Sci U S A. 2012;109:13698–703. doi:10.1073/pnas.1206625109.

82. Dores RM. Hagfish, genome duplications, and RFamide neuropeptide evolution. Endocrinology. 2011;152:4010–3. doi:10.1210/en.2011-1694.

83. Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F, et al. Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res. 2000;10:1890–902.

84. Catchen JM, Braasch I, Postlethwait JH. Conserved synteny and the zebrafish genome. Methods Cell Biol. 2011;104:259–85.

85. Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, et al. The spotted gar genome illuminates vertebrate evolution and facilitates human-to-teleost comparisons. Nat Genet. 2016;48:427–37. doi:10.1038/ng.3526.

86. Sato Y, Nishida M. Teleost fish with specific genome duplication as unique models of vertebrate evolution. Environ Biol Fish. 2010;88:169–88. doi:10.1007/s10641-010-9628-7.

87. Hiller M, Agarwal S, Notwell JH, Parikh R, Guturu H, Wenger AM, et al. Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: Application to zebrafish. Nucleic Acids Res. 2013;41:e151. doi:10.1093/nar/gkt557.

88. Wallace KN, Akhter S, Smith EM, Lorent K, Pack M. Intestinal growth and differentiation in zebrafish. Mech Dev. 2005;122:157–73.

89. Faro A, Boj SF, Ambrósio R, Broek O van den, Korving J, Clevers H. T-cell factor 4 (tcf7l2) is the main effector of wnt signaling during zebrafish intestine organogenesis. Zebrafish. 2009;6:59–68.

90. White R, Rose K, Zon L. Zebrafish cancer: The state of the art and the path forward. Nat Rev Cancer. 2013;13:624–36.

91. Parsons MJ, Pisharath H, Yusuff S, Moore JC, Siekmann AF, Lawson N, et al. Notch-responsive cells initiate the secondary transition in larval zebrafish pancreas. Mechanisms of Development. 2009;126:898–912. doi:10.1016/j.mod.2009.07.002.

92. Kinkel MD, Prince VE. On the diabetic menu: Zebrafish as a model for pancreas development and function. Bioessays. 2009;31:139–52.

93. Jurczyk A, Roy N, Bajwa R, Gut P, Lipson K, Yang C, et al. Dynamic glucoregulation and mammalian-like responses to metabolic and developmental disruption in zebrafish. General and Comparative Endocrinology. 2011;170:334–45. doi:10.1016/j.ygcen.2010.10.010.

94. Grada A, Weinbrecht K. Next-generation sequencing: Methodology and application. Journal of Investigative Dermatology. 2013;133:1–4. doi:10.1038/jid.2013.248.

95. Behjati S, Tarpey PS. What is next generation sequencing? Arch Dis Child Educ Pract Ed. 2013;98:236–8. doi:10.1136/archdischild-2013-304340.

96. Koch CM, Chiu SF, Akbarpour M, Bharat A, Ridge KM, Bartom ET, et al. A beginner's guide to analysis of RNA sequencing data. Am J Respir Cell Mol Biol. 2018;59:145–57. doi:10.1165/rcmb.2017-0430TR.

97. Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. The sanger FASTQ file format for sequences with quality scores, and the solexa/illumina FASTQ variants. Nucleic Acids Res. 2010;38:1767–71. doi:10.1093/nar/gkp1137.

98. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.

99. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nature Biotechnology. 2011;29:24–6. doi:10.1038/nbt.1754.

100. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92. doi:10.1093/bib/bbs017.

101. Variant review with the integrative genomics viewer cancer research. https://cancerres.aacrjournals.org/content/77/21/e31.long. Accessed 10 Jun 2020.

102. Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, et al. A comprehensive evaluation of normalization methods for illumina high-throughput RNA sequencing data analysis. Brief Bioinform. 2013;14:671–83. doi:10.1093/bib/bbs046.

103. Oshlack A, Wakefield MJ. Transcript length bias in RNA-seq data confounds systems biology. Biol Direct. 2009;4:14.

104. Ringnér M. What is principal component analysis? Nature Biotechnology. 2008;26:303–4. doi:10.1038/nbt0308-303.

105. Hauke J, Kossowski T. Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. Quaestiones Geographicae. 2011;30:87–93. doi:10.2478/v10117-011-0021-1.

106. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

107. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40. doi:10.1093/bioinformatics/btp616.

108. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40:4288–97. doi:10.1093/nar/gks042.

109. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 2013;41 Database issue:D377–386.

110. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48. doi:10.1186/1471-2105-10-48.

111. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLOS Computational Biology. 2007;3:e39. doi:10.1371/journal.pcbi.0030039.

112. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately down-regulated in human diabetes. Nature Genetics. 2003;34:267–73. doi:10.1038/ng1180.

113. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.

114. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.

115. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90–97.

116. Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung B, Smibert P, et al. Cell "hashing" with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. bioRxiv. 2017;237693. doi:10.1101/237693.

117. Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, et al. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nature Biotechnology. 2018;36:89–94. doi:10.1038/nbt.4042.

118. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Systems. 2019;8:329–337.e4. doi:10.1016/j.cels.2019.03.003.

119. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol.

2018;36:411–20. doi:10.1038/nbt.4096.

120. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. Cell. 2019;177:1888–1902.e21. doi:10.1016/j.cell.2019.05.031.

121. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biology. 2019;20:296. doi:10.1186/s13059-019-1874-1.

122. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2019;37:38–44. doi:10.1038/nbt.4314.

123. Maaten L van der, Hinton G. Visualizing data using t-SNE. Journal of Machine Learning Research. 2008;9 Nov:2579–605. http://www.jmlr.org/papers/v9/vandermaaten08a.html. Accessed 10 Jun 2020.

124. Maffei A, Liu Z, Witkowski P, Moschella F, Del Pozzo G, Liu E, et al. Identification of tissue-restricted transcripts in human islets. Endocrinology. 2004;145:4513–21.

125. Kutlu B, Burdick D, Baxter D, Rasschaert J, Flamez D, Eizirik DL, et al. Detailed transcriptome atlas of the pancreatic beta cell. BMC Med Genomics. 2009;2:3.

126. Szabat M, Luciani DS, Piret JM, Johnson JD. Maturation of adult beta-cells revealed using a pdx1/insulin dual-reporter lentivirus. Endocrinology. 2009;150:1627–35.

127. Marselli L, Thorne J, Dahiya S, Sgroi DC, Sharma A, Bonner-Weir S, et al. Gene expression profiles of beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. PLoS ONE. 2010;5:e11499.

128. Dorrell C, Schug J, Lin CF, Canaday PS, Fox AJ, Smirnova O, et al. Transcriptomes of the major human pancreatic cell types. Diabetologia. 2011;54:2832–44.

129. Morán I, Akerman I, Bunt M van de, Xie R, Benazra M, Nammo T, et al. Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab. 2012;16:435–48.

130. Nica AC, Ongen H, Irminger J-C, Bosco D, Berney T, Antonarakis SE, et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. Genome Res. 2013;23:1554–62.

131. Benner C, Meulen T van der, Cacéres E, Tigyi K, Donaldson CJ, Huising MO. The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. BMC Genomics. 2014;15:620.

132. Bramswig NC, Kaestner KH. Transcriptional and epigenetic regulation in human islets. Diabetologia. 2014;57:451–4.

133. Blodgett DM, Nowosielska A, Afik S, Pechhold S, Cura AJ, Kennedy NJ, et al. Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. Diabetes. 2015;64:3172–81.

134. Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Systems. 2016;3:346–360.e4. doi:10.1016/j.cels.2016.08.011.

135. Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, et al. A single-cell transcriptome atlas of the human pancreas. Cell Systems. 2016;3:385–394.e3. doi:10.1016/j.cels.2016.09.002.

136. Segerstolpe Å, Palasantza A, Eliasson P, Andersson E-M, Andréasson A-C, Sun X, et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. Cell Metab. 2016;24:593–607. doi:10.1016/j.cmet.2016.08.020.

137. Wang YJ, Schug J, Won K-J, Liu C, Naji A, Avrahami D, et al. Single-cell transcriptomics of the human endocrine pancreas. Diabetes. 2016;65:3028–38.

138. Lawlor N, George J, Bolisetty M, Kursawe R, Sun L, Sivakamasundari V, et al. Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. Genome Res. 2017;27:208–22.

139. Li J, Klughammer J, Farlik M, Penz T, Spittler A, Barbieux C, et al. Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell types. EMBO

Rep. 2016;17:178–87.

140. Adriaenssens AE, Svendsen B, Lam BYH, Yeo GSH, Holst JJ, Reimann F, et al. Transcriptomic profiling of pancreatic alpha, beta and delta cell populations identifies delta cells as a principal target for ghrelin in mouse islets. Diabetologia. 2016;59:2156–65.

141. Li L-C, Qiu W-L, Zhang Y-W, Xu Z-R, Xiao Y-N, Hou C, et al. Single-cell transcriptomic analyses reveal distinct dorsal/ventral pancreatic programs. EMBO Rep. 2018;19.

142. Murtaugh LC. Pancreas and beta-cell development: From the actual to the possible. Development. 2007;134:427–38.

143. Kimmel RA, Meyer D. Molecular regulation of pancreas development in zebrafish. Methods Cell Biol. 2010;100:261–80.

144. Verbruggen V, Ek O, Georlette D, Delporte F, Von Berg V, Detry N, et al. The pax6b homeodomain is dispensable for pancreatic endocrine cell differentiation in zebrafish. J Biol Chem. 2010;285:13863–73. doi:10.1074/jbc.M110.108019.

145. Delporte FM, Pasque V, Devos N, Manfroid I, Voz ML, Motte P, et al. Expression of zebrafish pax6b in pancreas is regulated by two enhancers containing highly conserved cis-elements bound by PDX1, PBX and PREP factors. BMC Developmental Biology. 2008;8:53. doi:10.1186/1471-213X-8-53.

146. Tarifeño-Saldivia E, Lavergne A, Bernard A, Padamata K, Bergemann D, Voz ML, et al. Transcriptome analysis of pancreatic cells across distant species highlights novel important regulator genes. BMC Biol. 2017;15. doi:10.1186/s12915-017-0362-x.

147. Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, Ninov N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR–cas9-induced genetic scars. Nature Biotechnology. 2018;36:469–73. doi:10.1038/nbt.4124.

148. Lu C-J, Fan X-Y, Guo Y-F, Cheng Z-C, Dong J, Chen J-Z, et al. Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development. J Mol Cell Biol. 2019;11:435–47.

149. Blum B, Hrvatin S, Schuetz C, Bonal C, Rezania A, Melton DA. Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3.

Nat Biotechnol. 2012;30:261–4. doi:10.1038/nbt.2141.

150. Jia S, Ivanov A, Blasevic D, Müller T, Purfürst B, Sun W, et al. Insm1 cooperates with neurod1 and foxa2 to maintain mature pancreatic β-cell function. EMBO J. 2015;34:1417–33.

151. Zhu Y, Liu Q, Zhou Z, Ikeda Y. PDX1, neurogenin-3, and MAFA: Critical transcription regulators for beta cell development and regeneration. Stem Cell Research & Therapy. 2017;8:240. doi:10.1186/s13287-017-0694-z.

152. Gao T, McKenna B, Li C, Reichert M, Nguyen J, Singh T, et al. Pdx1 maintains β-cell identity and function by repressing an α-cell program. Cell Metab. 2014;19:259–71. doi:10.1016/j.cmet.2013.12.002.

153. Fujimoto K, Polonsky KS. Pdx1 and other factors that regulate pancreatic beta-cell survival. Diabetes Obes Metab. 2009;11 Suppl 4:30–7.

154. Spaeth JM, Gupte M, Perelis M, Yang Y-P, Cyphert H, Guo S, et al. Defining a novel role for the pdx1 transcription factor in islet β-cell maturation and proliferation during weaning. Diabetes. 2017;66:2830–9. doi:10.2337/db16-1516.

155. Piccand J, Strasser P, Hodson DJ, Meunier A, Ye T, Keime C, et al. Rfx6 maintains the functional identity of adult pancreatic β cells. Cell Rep. 2014;9:2219–32.

156. Swisa A, Avrahami D, Eden N, Zhang J, Feleke E, Dahan T, et al. PAX6 maintains β cell identity by repressing genes of alternative islet cell types. J Clin Invest. 127:230–43. doi:10.1172/JCI88015.

157. Taylor BL, Liu F-F, Sander M. Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells. Cell Rep. 2013;4:1262–75. doi:10.1016/j.celrep.2013.08.010.

158. Ediger BN, Lim H-W, Juliana C, Groff DN, Williams LT, Dominguez G, et al. LIM domain–binding 1 maintains the terminally differentiated state of pancreatic β cells. J Clin Invest. 127:215–29. doi:10.1172/JCI88016.

159. Matschinsky FM, Ellerman JE. Metabolism of glucose in the islets of langerhans. J Biol Chem. 1968;243:2730–6. http://www.jbc.org/content/243/10/2730. Accessed 16 Dec 2019.

160. Hedeskov CJ. Mechanism of glucose-induced insulin secretion. Physiol Rev. 1980;60:442–509.

161. Campbell JE, Drucker DJ. Pharmacology, physiology, and mechanisms of incretin hormone action. Cell Metab. 2013;17:819–37.

162. Niemeyer H, Luz Cárdenas M de la, Rabajille E, Ureta T, Clark-Turri L, Peñaranda J. Sigmoidal kinetics of glucokinase. Enzyme. 1975;20:321–33.

163. Froguel P, Vaxillaire M, Sun F, Velho G, Zouali H, Butel MO, et al. Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. Nature. 1992;356:162–4.

164. Prentki M, Matschinsky FM, Madiraju SRM. Metabolic signaling in fuel-induced insulin secretion. Cell Metabolism. 2013;18:162–85. doi:10.1016/j.cmet.2013.05.018.

165. Wu H-H, Li Y-L, Liu N-J, Yang Z, Tao X-M, Du Y-P, et al. TCF7L2 regulates pancreatic β-cell function through PI3K/AKT signal pathway. Diabetol Metab Syndr. 2019;11. doi:10.1186/s13098-019-0449-3.

166. Heit JJ, Apelqvist AA, Gu X, Winslow MM, Neilson JR, Crabtree GR, et al. Calcineurin/NFAT signalling regulates pancreatic beta-cell growth and function. Nature. 2006;443:345–9.

167. Newsholme P, Cruzat V, Arfuso F, Keane K. Nutrient regulation of insulin secretion and action. J Endocrinol. 2014;221:R105–120.

168. Johnston NR, Mitchell RK, Haythorne E, Pessoa MP, Semplici F, Ferrer J, et al. Beta cell hubs dictate pancreatic islet responses to glucose. Cell Metab. 2016;24:389–401. doi:10.1016/j.cmet.2016.06.020.

169. Da Silva Xavier G, Rutter GA. Metabolic and functional heterogeneity in pancreatic β cells. J Mol Biol. 2019.

170. Singh SP, Janjuha S, Hartmann T, Kayisoglu Ö, Konantz J, Birke S, et al. Different developmental histories of beta-cells generate functional and proliferative heterogeneity during islet growth. Nat Commun. 2017;8:1–16. doi:10.1038/s41467-017-00461-3.

171. Salem V, Silva LD, Suba K, Georgiadou E, Gharavy SNM, Akhtar N, et al. Leader β-cells coordinate ca 2+ dynamics across pancreatic islets in vivo. Nat Metab. 2019;1:615–29. doi:10.1038/s42255-019-0075-2.

172. Thorel F, Damond N, Chera S, Wiederkehr A, Thorens B, Meda P, et al. Normal glucagon signaling and β-cell function after near-total α-cell ablation in adult mice. Diabetes. 2011;60:2872–82.

173. Thorel F, Népote V, Avril I, Kohno K, Desgraz R, Chera S, et al. Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature. 2010;464:1149–54.

174. Chera S, Baronnier D, Ghila L, Cigliola V, Jensen JN, Gu G, et al. Diabetes recovery by age-dependent conversion of pancreatic δ-cells into insulin producers. Nature. 2014;514:503–7.

175. Bramswig NC, Everett LJ, Schug J, Dorrell C, Liu C, Luo Y, et al. Epigenomic plasticity enables human pancreatic α to β cell reprogramming. J Clin Invest. 2013;123:1275–84.

176. Napolitano T, Avolio F, Courtney M, Vieira A, Druelle N, Ben-Othman N, et al. Pax4 acts as a key player in pancreas development and plasticity. Semin Cell Dev Biol. 2015;44:107–14.

177. Ben-Othman N, Vieira A, Courtney M, Record F, Gjernes E, Avolio F, et al. Long-term GABA administration induces alpha cell-mediated beta-like cell neogenesis. Cell. 2017;168:73–85.e11.

178. Unger RH, Orci L. The essential role of glucagon in the pathogenesis of diabetes mellitus. Lancet. 1975;1:14–6.

179. Gerich JE. Lilly lecture 1988. Glucose counterregulation and its impact on diabetes mellitus. Diabetes. 1988;37:1608–17.

180. Cryer PE. Minireview: Glucagon in the pathogenesis of hypoglycemia and hyperglycemia in diabetes. Endocrinology. 2012;153:1039–48. doi:10.1210/en.2011-1499.

181. Dunning BE, Gerich JE. The role of alpha-cell dysregulation in fasting and postprandial hyperglycemia in type 2 diabetes and therapeutic implications. Endocr Rev. 2007;28:253–83.

182. Gromada J, Franklin I, Wollheim CB. Alpha-cells of the endocrine pancreas: 35 years of research but the enigma remains. Endocr Rev. 2007;28:84–116.

183. Rouillé Y, Westermark G, Martin SK, Steiner DF. Proglucagon is processed to glucagon by prohormone convertase PC2 in alpha TC1-6 cells. Proc Natl Acad Sci USA. 1994;91:3242–6.

184. Hinke SA, Pospisilik JA, Demuth HU, Mannhart S, Kühn-Wache K, Hoffmann T, et al. Dipeptidyl peptidase IV (DPIV/CD26) degradation of glucagon. Characterization of glucagon degradation products and DPIV-resistant analogs. J Biol Chem. 2000;275:3827–34.

185. Zhang J, McKenna LB, Bogue CW, Kaestner KH. The diabetes gene hhex maintains δ-cell differentiation and islet function. Genes Dev. 2014;28:829–34. doi:10.1101/gad.235499.113.

186. Stefan Y, Orci L, Malaisse-Lagae F, Perrelet A, Patel Y, Unger RH. Quantitation of endocrine cell content in the pancreas of nondiabetic and diabetic humans. Diabetes. 1982;31:694–700.

187. Khandekar N, Berning BA, Sainsbury A, Lin S. The role of pancreatic polypeptide in the regulation of energy homeostasis. Molecular and Cellular Endocrinology. 2015;418:33–41. doi:10.1016/j.mce.2015.06.028.

188. Schwartz TW. Pancreatic polypeptide: A hormone under vagal control. Gastroenterology. 1983;85:1411–25.

189. Holzer P, Reichmann F, Farzi A. Neuropeptide y, peptide YY and pancreatic polypeptide in the gut–brain axis. Neuropeptides. 2012;46:261–74. doi:10.1016/j.npep.2012.08.005.

190. Wierup N, Sundler F, Heller RS. The islet ghrelin cell. J Mol Endocrinol. 2014;52:R35–49.

191. Broglio F, Arvat E, Benso A, Gottero C, Muccioli G, Papotti M, et al. Ghrelin, a natural GH secretagogue produced by the stomach, induces hyperglycemia and reduces insulin secretion in humans. J Clin Endocrinol Metab. 2001;86:5083–6.

192. Broglio F, Gottero C, Benso A, Prodam F, Destefanis S, Gauna C, et al. Effects of

ghrelin on the insulin and glycemic responses to glucose, arginine, or free fatty acids load in humans. J Clin Endocrinol Metab. 2003;88:4268–72.

193. Egido EM, Rodriguez-Gallardo J, Silvestre RA, Marco J. Inhibitory effect of ghrelin on insulin and pancreatic somatostatin secretion. Eur J Endocrinol. 2002;146:241–4.

194. Mitchell RK, Nguyen-Tu M-S, Chabosseau P, Callingham RM, Pullen TJ, Cheung R, et al. The transcription factor pax6 is required for pancreatic β cell identity, glucose-regulated ATP synthesis, and ca2+ dynamics in adult mice. J Biol Chem. 2017;292:8892–906. doi:10.1074/jbc.M117.784629.

195. Biemar F, Argenton F, Schmidtke R, Epperlein S, Peers B, Driever W. Pancreas development in zebrafish: Early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. Dev Biol. 2001;230:189–203.

196. Field HA, Dong PDS, Beis D, Stainier DYR. Formation of the digestive system in zebrafish. II. Pancreas morphogenesis. Dev Biol. 2003;261:197–208.

197. Wendik B, Maier E, Meyer D. Zebrafish mnx genes in endocrine and exocrine pancreas formation. Developmental Biology. 2004;268:372–83. doi:10.1016/j.ydbio.2003.12.026.

198. Zecchin E, Mavropoulos A, Devos N, Filippi A, Tiso N, Meyer D, et al. Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates. Developmental Biology. 2004;268:174–84. doi:10.1016/j.ydbio.2003.12.016.

199. Kimmel CB, Warga RM, Schilling TF. Origin and organization of the zebrafish fate map. Development. 1990;108:581–94.

200. Warga RM, Kimmel CB. Cell movements during epiboly and gastrulation in zebrafish. Development. 1990;108:569–80.

201. Dougan ST, Warga RM, Kane DA, Schier AF, Talbot WS. The role of the zebrafish nodal-related genes squint and cyclops in patterning of mesendoderm. Development. 2003;130:1837–51.

202. Aoki TO, Mathieu J, Saint-Etienne L, Rebagliati MR, Peyriéras N, Rosa FM. Regulation of nodal signalling and mesendoderm formation by TARAM-a, a TGFbeta-related type i receptor. Dev Biol. 2002;241:273–88.

203. Reiter JF, Kikuchi Y, Stainier DY. Multiple roles for gata5 in zebrafish endoderm formation. Development. 2001;128:125–35.

204. Poulain M, Lepage T. Mezzo, a paired-like homeobox protein is an immediate target of nodal signalling and regulates endoderm specification in zebrafish. Development. 2002;129:4901–14.

205. Aoki TO, David NB, Minchiotti G, Saint-Etienne L, Dickmeis T, Persico GM, et al. Molecular integration of casanova in the nodal signalling pathway controlling endoderm formation. Development. 2002;129:275–86.

206. Odenthal J, Nüsslein-Volhard C. Fork head domain genes in zebrafish. Dev Genes Evol. 1998;208:245–58.

207. Wallace KN, Pack M. Unique and conserved aspects of gut development in zebrafish. Dev Biol. 2003;255:12–29.

208. McLin VA, Rankin SA, Zorn AM. Repression of wnt/beta-catenin signaling in the anterior endoderm is essential for liver and pancreas development. Development. 2007;134:2207–17.

209. Manfroid I, Delporte F, Baudhuin A, Motte P, Neumann CJ, Voz ML, et al. Reciprocal endoderm-mesoderm interactions mediated by fgf24 and fgf10 govern pancreas development. Development. 2007;134:4011–21.

210. Tiso N, Filippi A, Pauls S, Bortolussi M, Argenton F. BMP signalling regulates anteroposterior endoderm patterning in zebrafish. Mech Dev. 2002;118:29–37.

211. Song J, Kim HJ, Gong Z, Liu N-A, Lin S. Vhnf1 acts downstream of bmp, fgf, and RA signals to regulate endocrine beta cell development in zebrafish. Dev Biol. 2007;303:561–75.

212. Lokmane L, Haumaitre C, Garcia-Villalba P, Anselme I, Schneider-Maunoury S, Cereghini S. Crucial role of vHNF1 in vertebrate hepatic specification. Development. 2008;135:2777–86.

213. Stafford D, White RJ, Kinkel MD, Linville A, Schilling TF, Prince VE. Retinoids signal directly to zebrafish endoderm to specify insulin-expressing beta-cells. Development.

2006;133:949–56.

214. Tehrani Z, Lin S. Antagonistic interactions of hedgehog, bmp and retinoic acid signals control zebrafish endocrine pancreas development. Development. 2011;138:631–40.

215. Tehrani Z, Lin S. Endocrine pancreas development in zebrafish. Cell Cycle. 2011;10:3466–72.

216. Chung W-S, Stainier DYR. Intra-endodermal interactions are required for pancreatic beta cell induction. Dev Cell. 2008;14:582–93.

217. Warga RM, Nüsslein-Volhard C. Origin and development of the zebrafish endoderm. Development. 1999;126:827–38.

218. Kikuchi Y, Verkade H, Reiter JF, Kim C-H, Chitnis AB, Kuroiwa A, et al. Notch signaling can regulate endoderm formation in zebrafish. Dev Dyn. 2004;229:756–62.

219. Poulain M, Fürthauer M, Thisse B, Thisse C, Lepage T. Zebrafish endoderm formation is regulated by combinatorial nodal, FGF and BMP signalling. Development. 2006;133:2189–200.

220. diIorio P, Alexa K, Choe S-K, Etheridge L, Sagerström CG. TALE-family homeodomain proteins regulate endodermal sonic hedgehog expression and pattern the anterior endoderm. Dev Biol. 2007;304:221–31. doi:10.1016/j.ydbio.2006.12.024.

221. Noël ES, Casal-Sueiro A, Busch-Nentwich E, Verkade H, Dong PDS, Stemple DL, et al. Organ-specific requirements for hdac1 in liver and pancreas formation. Dev Biol. 2008;322:237–50. doi:10.1016/j.ydbio.2008.06.040.

222. Kinkel MD, Eames SC, Alonzo MR, Prince VE. Cdx4 is required in the endoderm to localize the pancreas and limit beta-cell number. Development. 2008;135:919–29.

223. Tiso N, Moro E, Argenton F. Zebrafish pancreas development. Molecular and Cellular Endocrinology. 2009;312:24–30. doi:10.1016/j.mce.2009.04.018.

224. Dong PDS, Munson CA, Norton W, Crosnier C, Pan X, Gong Z, et al. Fgf10 regulates hepatopancreatic ductal system patterning and differentiation. Nat Genet. 2007;39:397–402.

225. Zecchin E, Filippi A, Biemar F, Tiso N, Pauls S, Ellertsdottir E, et al. Distinct delta and jagged genes control sequential segregation of pancreatic cell types from precursor pools in zebrafish. Dev Biol. 2007;301:192–204.

226. Puri S, Hebrok M. Cellular plasticity within the pancreas–lessons learned from development. Dev Cell. 2010;18:342–56.

227. Ahlgren U, Jonsson J, Jonsson L, Simu K, Edlund H. Beta-cell-specific inactivation of the mouse ipf1/pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes Dev. 1998;12:1763–8.

228. Ashizawa S, Brunicardi FC, Wang X-P. PDX-1 and the pancreas. Pancreas. 2004;28:109–20.

229. Wang YJ, Park JT, Parsons MJ, Leach SD. Fate mapping of ptf1a-expressing cells during pancreatic organogenesis and regeneration in zebrafish. Dev Dyn. 2015;244:724–35. doi:10.1002/dvdy.24271.

230. Hoang CQ, Hale MA, Azevedo-Pouly AC, Elsässer HP, Deering TG, Willet SG, et al. Transcriptional maintenance of pancreatic acinar identity, differentiation, and homeostasis by PTF1A. Mol Cell Biol. 2016;36:3033–47. doi:10.1128/MCB.00358-16.

231. Flasse LC, Pirson JL, Stern DG, Von Berg V, Manfroid I, Peers B, et al. Ascl1b and neurod1, instead of neurog3, control pancreatic endocrine cell fate in zebrafish. BMC Biol. 2013;11:78.

232. Gradwohl G, Dierich A, LeMeur M, Guillemot F. Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. Proc Natl Acad Sci USA. 2000;97:1607–11.

233. Schwitzgebel VM, Scheel DW, Conners JR, Kalamaras J, Lee JE, Anderson DJ, et al. Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. Development. 2000;127:3533–42.

234. Gu G, Dubauskaite J, Melton DA. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development. 2002;129:2447–57.

235. Ghaye AP, Bergemann D, Tarifeño-Saldivia E, Flasse LC, Von Berg V, Peers B, et al. Progenitor potential of nkx6.1-expressing cells throughout zebrafish life and during beta cell regeneration. BMC Biol. 2015;13:70.

236. Schaffer AE, Freude KK, Nelson SB, Sander M. Ptf1a and nkx6 transcription factors function as antagonistic lineage determinants in multipotent pancreatic progenitors. Dev Cell. 2010;18:1022–9. doi:10.1016/j.devcel.2010.05.015.

237. Miralles F, Czernichow P, Scharfmann R. Follistatin regulates the relative proportions of endocrine versus exocrine tissue during pancreatic development. Development. 1998;125:1017–24.

238. Duvillié B, Attali M, Bounacer A, Ravassard P, Basmaciogullari A, Scharfmann R. The mesenchyme controls the timing of pancreatic beta-cell differentiation. Diabetes. 2006;55:582–9.

239. Apelqvist A, Li H, Sommer L, Beatus P, Anderson DJ, Honjo T, et al. Notch signalling controls pancreatic cell differentiation. Nature. 1999;400:877–81.

240. Esni F, Ghosh B, Biankin AV, Lin JW, Albert MA, Yu X, et al. Notch inhibits ptf1 function and acinar cell differentiation in developing mouse and zebrafish pancreas. Development. 2004;131:4213–24.

241. Benitez CM, Goodyer WR, Kim SK. Deconstructing pancreas developmental biology. Cold Spring Harb Perspect Biol. 2012;4. doi:10.1101/cshperspect.a012401.

242. Zhang H, Ables ET, Pope CF, Washington MK, Hipkens S, Means AL, et al. Multiple, temporal-specific roles for HNF6 in pancreatic endocrine and ductal differentiation. Mech Dev. 2009;126:958–73.

243. Burstin J von, Reichert M, Wescott MP, Rustgi AK. The pancreatic and duodenal homeobox protein PDX-1 regulates the ductal specific keratin 19 through the degradation of MEIS1 and DNA binding. PLoS ONE. 2010;5:e12311.

244. Manfroid I, Ghaye A, Naye F, Detry N, Palm S, Pan L, et al. Zebrafish sox9b is crucial for hepatopancreatic duct development and pancreatic endocrine cell regeneration. Dev Biol. 2012;366:268–78.

245. Zhang D, Gates KP, Barske L, Wang G, Lancman JJ, Zeng X-XI, et al. Endoderm jagged induces liver and pancreas duct lineage in zebrafish. Nat Commun. 2017;8:769.

246. Mavropoulos A, Devos N, Biemar F, Zecchin E, Argenton F, Edlund H, et al. Sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. Dev Biol. 2005;285:211–23.

247. Binot A-C, Manfroid I, Flasse L, Winandy M, Motte P, Martial JA, et al. Nkx6.1 and nkx6.2 regulate alpha- and beta-cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. Dev Biol. 2010;340:397–407.

248. Djiotsa J, Verbruggen V, Giacomotto J, Ishibashi M, Manning E, Rinkwitz S, et al. Pax4 is not essential for beta-cell differentiation in zebrafish embryos but modulates alpha-cell generation by repressing arx gene expression. BMC Dev Biol. 2012;12:37.

249. Soyer J, Flasse L, Raffelsberger W, Beucher A, Orvain C, Peers B, et al. Rfx6 is an ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. Development. 2010;137:203–12.

250. Arkhipova V, Wendik B, Devos N, Ek O, Peers B, Meyer D. Characterization and regulation of the hb9/mnx1 beta-cell progenitor specific enhancer in zebrafish. Dev Biol. 2012;365:290–302.

251. Dalgin G, Ward AB, Hao LT, Beattie CE, Nechiporuk A, Prince VE. Zebrafish mnx1 controls cell fate choice in the developing endocrine pancreas. Development. 2011;138:4597–608. doi:10.1242/dev.067736.

252. Kimmel RA, Onder L, Wilfinger A, Ellertsdottir E, Meyer D. Requirement for pdx1 in specification of latent endocrine progenitors in zebrafish. BMC Biol. 2011;9:75.

253. Sosa-Pineda B. The gene pax4 is an essential regulator of pancreatic beta-cell development. Mol Cells. 2004;18:289–94.

254. Wang J, Elghazi L, Parker SE, Kizilocak H, Asano M, Sussel L, et al. The concerted activities of pax4 and nkx2.2 are essential to initiate pancreatic beta-cell differentiation. Dev Biol. 2004;266:178–89.

255. Collombat P, Hecksher-Sørensen J, Serup P, Mansouri A. Specifying pancreatic endocrine cell fates. Mechanisms of Development. 2006;123:501–12.

doi:10.1016/j.mod.2006.05.006.

256. Collombat P, Xu X, Ravassard P, Sosa-Pineda B, Dussaud S, Billestrup N, et al. The ectopic expression of pax4 in the mouse pancreas converts progenitor cells into alpha and subsequently beta cells. Cell. 2009;138:449–62.

257. Kordowich S, Collombat P, Mansouri A, Serup P. Arx and nkx2.2 compound deficiency redirects pancreatic alpha- and beta-cell differentiation to a somatostatin/ghrelin co-expressing cell lineage. BMC Dev Biol. 2011;11:52.

258. Sussel L, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL, et al. Mice lacking the homeodomain transcription factor nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. Development. 1998;125:2213–21.

259. Prado CL, Pugh-Bernard AE, Elghazi L, Sosa-Pineda B, Sussel L. Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development. Proc Natl Acad Sci USA. 2004;101:2924–9.

260. Pauls S, Zecchin E, Tiso N, Bortolussi M, Argenton F. Function and regulation of zebrafish nkx2.2a during development of pancreatic islet and ducts. Dev Biol. 2007;304:875–90.

261. Wilfinger A, Arkhipova V, Meyer D. Cell type and tissue specific function of islet genes in zebrafish pancreas development. Dev Biol. 2013;378:25–37.

262. Villasenor A, Gauvrit S, Collins MM, Maischein H-M, Stainier DYR. Hhex regulates the specification and growth of the hepatopancreatic ductal system. Developmental Biology. 2020;458:228–36. doi:10.1016/j.ydbio.2019.10.021.

263. Cvekl A, Callaerts P. PAX6: 25th anniversary and more to learn. Experimental Eye Research. 2017;156:10–21. doi:10.1016/j.exer.2016.04.017.

264. Callaerts P, Halder G, Gehring WJ. PAX-6 in development and evolution. Annu Rev Neurosci. 1997;20:483–532.

265. Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, Heyningen V van. Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. Hum Mol Genet.

2001;10:2049–59.

266. Kleinjan DA, Seawright A, Mella S, Carr CB, Tyas DA, Simpson TI, et al. Long-range downstream enhancers are essential for pax6 expression. Dev Biol. 2006;299:563–81.

267. Fabian P, Kozmikova I, Kozmik Z, Pantzartzi CN. Pax2/5/8 and pax6 alternative splicing events in basal chordates and vertebrates: A focus on paired box domain. Front Genet. 2015;6:228.

268. Vance KW, Sansom SN, Lee S, Chalei V, Kong L, Cooper SE, et al. The long non-coding RNA paupar regulates the expression of both local and distal genes. EMBO J. 2014;33:296–311.

269. Nornes S, Clarkson M, Mikkola I, Pedersen M, Bardsley A, Martinez JP, et al. Zebrafish contains two pax6 genes involved in eye development. Mech Dev. 1998;77:185–96.

270. Kleinjan DA, Bancewicz RM, Gautier P, Dahm R, Schonthaler HB, Damante G, et al. Subfunctionalization of duplicated zebrafish pax6 genes by cis-regulatory divergence. PLoS Genet. 2008;4:e29.

271. Hickmott JW, Chen C-Y, Arenillas DJ, Korecki AJ, Lam SL, Molday LL, et al. PAX6 MiniPromoters drive restricted expression from rAAV in the adult mouse retina. Mol Ther Methods Clin Dev. 2016;3:16051.

272. Kammandel B, Chowdhury K, Stoykova A, Aparicio S, Brenner S, Gruss P. Distinct cis-essential modules direct the time-space pattern of the pax6 gene activity. Dev Biol. 1999;205:79–97.

273. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the pax6 paired domain are regulated by alternative splicing. Genes Dev. 1994;8:2022–34.

274. Kozmik Z, Czerny T, Busslinger M. Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of pax6 and pax8. EMBO J. 1997;16:6793–803. doi:10.1093/emboj/16.22.6793.

275. Haubst N, Berger J, Radjendirane V, Graw J, Favor J, Saunders GF, et al. Molecular dissection of pax6 function: The specific roles of the paired domain and homeodomain in

brain development. Development. 2004;131:6131–40.

276. Sun J, Rockowitz S, Xie Q, Ashery-Padan R, Zheng D, Cvekl A. Identification of in vivo DNA-binding mechanisms of pax6 and reconstruction of pax6-dependent gene regulatory networks during forebrain and lens development. Nucleic Acids Res. 2015;43:6827–46. doi:10.1093/nar/gkv589.

277. Sander M, Neubüser A, Kalamaras J, Ee HC, Martin GR, German MS. Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes Dev. 1997;11:1662–73. doi:10.1101/gad.11.13.1662.

278. St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, Gruss P. Pax6 is required for differentiation of glucagon-producing α-cells in mouse pancreas. Nature. 1997;387:406–9. doi:10.1038/387406a0.

279. Hart AW, Mella S, Mendrychowski J, Heyningen V van, Kleinjan DA. The developmental regulator pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. PLoS One. 2013;8. doi:10.1371/journal.pone.0054173.

280. Heller RS, Jenny M, Collombat P, Mansouri A, Tomasetto C, Madsen OD, et al. Genetic determinants of pancreatic epsilon-cell development. Dev Biol. 2005;286:217–24.

281. Gosmain Y, Marthinet E, Cheyssac C, Guérardel A, Mamin A, Katz LS, et al. Pax6 controls the expression of critical genes involved in pancreatic {alpha} cell differentiation and function. J Biol Chem. 2010;285:33381–93.

282. Katz LS, Gosmain Y, Marthinet E, Philippe J. Pax6 regulates the proglucagon processing enzyme PC2 and its chaperone 7B2. Mol Cell Biol. 2009;29:2322–34. doi:10.1128/MCB.01543-08.

283. Gosmain Y, Katz LS, Masson MH, Cheyssac C, Poisson C, Philippe J. Pax6 is crucial for β-cell function, insulin biosynthesis, and glucose-induced insulin secretion. Mol Endocrinol. 2012;26:696–709. doi:10.1210/me.2011-1256.

284. Yasuda T, Kajimoto Y, Fujitani Y, Watada H, Yamamoto S, Watarai T, et al. PAX6 mutation as a genetic factor common to aniridia and glucose intolerance. Diabetes. 2002;51:224–30. doi:10.2337/diabetes.51.1.224.

285. Nishi M, Sasahara M, Shono T, Saika S, Yamamoto Y, Ohkawa K, et al. A case of novel de novo paired box gene 6 (PAX6) mutation with early-onset diabetes mellitus and aniridia. Diabet Med. 2005;22:641–4.

286. Pagliuca FW, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, et al. Generation of functional human pancreatic β cells in vitro. Cell. 2014;159:428–39.

287. Millman JR, Xie C, Van Dervort A, Gürtler M, Pagliuca FW, Melton DA. Generation of stem cell-derived β-cells from patients with type 1 diabetes. Nat Commun. 2016;7:11463.

288. Melton DA. Applied developmental biology: Making human pancreatic beta cells for diabetics. Curr Top Dev Biol. 2016;117:65–73.

289. Tritschler S, Theis FJ, Lickert H, Böttcher A. Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas. Mol Metab. 2017;6:974–90.

290. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13:397–406. doi:10.1074/mcp.M113.035600.

291. Holmstrom SR, Deering T, Swift GH, Poelwijk FJ, Mangelsdorf DJ, Kliewer SA, et al. LRH-1 and PTF1-l coregulate an exocrine pancreas-specific transcriptional network for digestive function. Genes Dev. 2011;25:1674–9.

292. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using smart-seq2. Nat Protoc. 2014;9:171–81. doi:10.1038/nprot.2014.006.

293. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

294. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. Mol Cell. 2010;38:576–89.

295. Daley T, Smith AD. Predicting the molecular complexity of sequencing libraries. Nat Methods. 2013;10:325–7. doi:10.1038/nmeth.2375.

296. Gutierrez JA, Solenberg PJ, Perkins DR, Willency JA, Knierman MD, Jin Z, et al. Ghrelin octanoylation mediated by an orphan lipid transferase. Proc Natl Acad Sci USA. 2008;105:6320–5.

297. Al Massadi O, Tschöp MH, Tong J. Ghrelin acylation and metabolic control. Peptides. 2011;32:2301–8. doi:10.1016/j.peptides.2011.08.020.

298. Leonard J, Peers B, Johnson T, Ferreri K, Lee S, Montminy MR. Characterization of somatostatin transactivating factor-1, a novel homeobox factor that stimulates somatostatin expression in pancreatic islet cells. Mol Endocrinol. 1993;7:1275–83.

299. Rouillé Y, Bianchi M, Irminger JC, Halban PA. Role of the prohormone convertase PC2 in the processing of proglucagon to glucagon. FEBS Lett. 1997;413:119–23.

300. Buckland B. Role of CaMKII in pancreatic alpha cells. Electronic Theses and Dissertations. 2013. https://digitalcommons.du.edu/etd/96.

301. Hu C, Jia W, Zhang R, Wang C, Lu J, Wu H, et al. Effect of RBP4 gene variants on circulating RBP4 concentration and type 2 diabetes in a chinese population. Diabet Med. 2008;25:11–8.

302. Kovacs P, Geyer M, Berndt J, Klöting N, Graham TE, Böttcher Y, et al. Effects of genetic variation in the human retinol binding protein-4 gene (RBP4) on insulin resistance and fat depot-specific mRNA expression. Diabetes. 2007;56:3095–100.

303. Munkhtulga L, Nagashima S, Nakayama K, Utsumi N, Yanagisawa Y, Gotoh T, et al. Regulatory SNP in the RBP4 gene modified the expression in adipocytes and associated with BMI. Obesity (Silver Spring). 2010;18:1006–14.

304. Zito E, Chin K-T, Blais J, Harding HP, Ron D. ERO1-beta, a pancreas-specific disulfide oxidase, promotes insulin biogenesis and glucose homeostasis. J Cell Biol. 2010;188:821–32.

305. Askwith T, Zeng W, Eggo MC, Stevens MJ. Oxidative stress and dysregulation of the taurine transporter in high-glucose-exposed human schwann cells: Implications for pathogenesis of diabetic neuropathy. Am J Physiol Endocrinol Metab. 2009;297:E620–628.

306. El Idrissi A, El Hilali F, Rotondo S, Sidime F. Effects of taurine supplementation on

neuronal excitability and glucose homeostasis. Adv Exp Med Biol. 2017;975 Pt 1:271–9.

307. Taneera J, Fadista J, Ahlqvist E, Atac D, Ottosson-Laakso E, Wollheim CB, et al. Identification of novel genes for glucose metabolism based upon expression pattern in human islets and effect on insulin secretion and glycemia. Hum Mol Genet. 2015;24:1945–55.

308. Gage BK, Asadi A, Baker RK, Webber TD, Wang R, Itoh M, et al. The role of ARX in human pancreatic endocrine specification. PLoS ONE. 2015;10:e0144100.

309. Petri A, Ahnfelt-Rønne J, Frederiksen KS, Edwards DG, Madsen D, Serup P, et al. The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice. J Mol Endocrinol. 2006;37:301–16.

310. Kulkarni SS, Salehzadeh F, Fritz T, Zierath JR, Krook A, Osler ME. Mitochondrial regulators of fatty acid metabolism reflect metabolic dysfunction in type 2 diabetes mellitus. Metab Clin Exp. 2012;61:175–85.

311. Jacobsson B, Carlström A, Platz A, Collins VP. Transthyretin messenger ribonucleic acid expression in the pancreas and in endocrine tumors of the pancreas and gut. J Clin Endocrinol Metab. 1990;71:875–80.

312. Ruiz de Azua I, Scarselli M, Rosemond E, Gautam D, Jou W, Gavrilova O, et al. RGS4 is a negative regulator of insulin release from pancreatic beta-cells in vitro and in vivo. Proc Natl Acad Sci USA. 2010;107:7999–8004.

313. Epstein J, Cai J, Glaser T, Jepeal L, Maas R. Identification of a pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. J Biol Chem. 1994;269:8355–61.

314. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017;6.